

**THE REGULATION OF YOLK PROTEIN 3 GENE  
EXPRESSION IN *DROSOPHILA MELANOGASTER***

ELAINE RONALDSON

Thesis presented for the Degree of Doctor of Philosophy  
University of Edinburgh  
1995



## ACKNOWLEDGEMENTS

I would firstly like to thank my supervisor, Mary Bownes, for her encouragement and advice throughout the time I have spent in her group.

I would also like to thank past and present members of the group for their help and support throughout the duration of my stay in the Bownes group, especially:

- Claire, Kathleen, Nian, Susan and Brenda in the early days and Alberto and Roger for their entertaining 'conversation'
- Debbie, Neil, Colin, Lucy, Bryce, Angela, Janis, Craig and Debiao for the good working atmosphere
- Carol McClean for pointing me in the right direction and for expert microinjection advice
- All the media and washing up team for helping things to run smoothly
- Isobel Black, Pauline Morgan and Jean Milne for last-minute typing of papers and grant proposals
- Graham Brown and Sir Frank Johnstone for their excellent last-minute photographic work
  
- Many thanks to Roger, Brenda, Craig and Alan for proofreading

I would also like to thank many friends who made Edinburgh an enjoyable experience, especially Marion, Richard, Chris and Shirley, the staff and locals of 'Preservation Hall', the staff and locals of 'The Holyrood Tavern' and the staff and locals of 'The Baron Bailie'.

Finally, I thank my parents, Helen and Robert, for the opportunity in the first place and for their continuous love and support. That includes you too Scott.

## **DECLARATION**

I declare that this thesis was composed by myself, and that the work described is my own, unless otherwise stated.

## ABSTRACT

In *Drosophila*, the three *yolk protein* (*yp*) genes are transcribed in a sex-, tissue- and developmentally specific manner, providing an ideal system in which to investigate the factors involved in their regulation. The yolk proteins are synthesised in the fat body of adult females, and in the ovarian follicle cells surrounding the developing oocyte during stages 8-10 of oogenesis. The mechanisms governing *yp* transcription are therefore expected to involve sex- and tissue-specific factors that are present at the correct developmental stages for gene activation. This project involved an analysis of the *yolk protein 3* (*yp3*) gene and its flanking sequences in order to locate *cis*-acting DNA sequences necessary for the highly specific pattern of expression which, in the longer term, could be used to isolate the proteins interacting with them. Using P-element mediated germ-line transformation, it was demonstrated that a 747 bp promoter region was sufficient to direct sex-specific expression in the female fat body and both the temporal- and cell-type-specificity of expression during oogenesis. Further analysis of this region separated two elements capable of independently governing *yp3* transcription in these tissues and a sub-fragment from the region conferring ovarian expression was used in gel retardation assays to demonstrate the presence of putative *trans*-acting proteins in ovarian nuclear extracts. No other sequences in the upstream, downstream or coding regions were identified that were autonomously involved in *yp3* expression.

Since *yolk protein* expression was known to be under nutritional and hormonal regulation, the *in vivo* influence of the steroid hormones, ecdysone and juvenile hormone on isolated regions of *yp3* was also investigated. Finally, because the *yolk protein* genes are coordinately transcribed, the mechanisms and factors involved are expected to be the same for all three genes, and the results of this research are thus compared with those reported for *yp1* and *yp2* in an attempt to clarify current knowledge on the transcriptional control of this small gene family.



## ABBREVIATIONS

Amp	Ampicillin
ATP	Adenosine-5'-triphosphate
bp	Base pair
BSA	Bovine serum albumin
°C	Degrees centigrade
cDNA	Complementary deoxyribonucleic acid
Ci	Curies
cm	Centimetre(s)
cpm	Counts per minute
dATP	Deoxyadenosine-5'-triphosphate
dCTP	Deoxycytosine-5'-triphosphate
dGTP	Deoxyguanosine-5'-triphosphate
dTTP	Deoxythymidine-5'-triphosphate
dNTP	Deoxynucleotide-5'-triphosphate
dH <sub>2</sub> O	Distilled water
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
Dnase	Deoxyribonuclease
DTT	Dithiothreitol
EDTA	Diaminoethanetetra-acetic acid
g	Gram(s)
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
kb	Kilobase
kD	KiloDalton
Klenow	Large fragment of DNA polymerase I
Krpm	Kilorevolutions per minute
M	Molar

μCi	Microcurie(s)
μg	Microgram(s)
μl	Microlitre(s)
μM	Micromolar
mA	Milliampere(s)
mCi	Millicurie(s)
mg	Milligram(s)
ml	Millilitre(s)
mm	Millimetre(s)
mM	Millimolar
min	Minutes
m wt	Molecular weight
ng	Nanogram(s)
nmol	Nanomole(s)
OLB	Oligo labelling buffer
OD	Optical density
[ <sup>32</sup> P]	β-emitting isotope of phosphorous
PEG	Polyethylene glycol
%	Percentage
pg	Picogram(s)
pH	-log <sup>10</sup> (Hydrogen ion concentration)
psi	Pounds per square inch
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
R.T.	Room temperature
[ <sup>35</sup> S]	β-emitting isotope of sulphur
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

Temed	NNN'N'-tetra-methyl-1,2-diamino-ethane
Tris	Tris(hydroxymethyl)-amino-methane
Triton-X-100	Octylphenoxypolyethoxyethanol
U	Units
UV	Ultraviolet
V	Volts
v/v	Volume per volume
w/v	Weight per volume
<i>yp1</i>	Yolk protein gene 1
<i>yp2</i>	Yolk protein gene 2
<i>yp3</i>	Yolk protein gene 3
YP1	Yolk Protein 1
YP2	Yolk Protein 2
YP3	Yolk Protein 3

# CONTENTS

Title page	i
Acknowledgements	ii
Declaration	iii
Abstract	iv
Abbreviations	v
Contents	viii

## CHAPTER 1 - Introduction

### - The Yolk Proteins And Their Genes in *Drosophila melanogaster*

1.1	Oogenesis	1
1.2	Vitellogenesis	3
1.3	Biochemical Analysis of the Yolk Proteins	4
1.4	Factors Affecting YP Synthesis and Vitellogenesis	7
1.4.1	Nutrition	7
1.4.2	Hormones	8
1.4.2.1	Juvenile Hormone	8
1.4.2.2	20-Hydroxyecdysone	11
1.4.3	Sex-determination Genes	12
1.4.3.1	Genetic Studies	15
1.4.3.2	Molecular Studies	17
1.4.3.3	<i>doublesex</i> and the <i>yolk protein</i> genes	18
1.5	Molecular Analysis of the <i>yolk protein</i> Genes	22
1.5.1	<i>yp1</i> and <i>yp2</i> <i>Cis</i> -acting Sequences	23
1.5.2	Identification of Fat Body <i>Cis</i> -acting Elements Regulating <i>yp1</i> and <i>yp2</i> Gene Expression	24
1.5.3	<i>Trans</i> -acting Factors that Bind to the <i>yp1/yp2</i> Fat Body Enhancer	26
1.5.4	Identification of Ovarian <i>Cis</i> -acting Elements Regulating <i>yp1</i> and <i>yp2</i> Gene Expression	28

1.5.5	Ovarian <i>Trans</i> -acting Factors	31
1.5.6	<i>yp3</i> <i>Cis</i> -acting Sequences	34
1.5.7	<i>yp3</i> <i>Trans</i> -acting Factors	36
CHAPTER 2 - Materials and Methods		37
2.1	Media	37
2.1.1	Bacterial Media	37
2.1.2	<i>Drosophila</i> Media	38
2.2	Bacterial, Plasmid and <i>Drosophila</i> Stocks	38
2.3	Solutions	41
2.3.1	Sequencing Gel Solutions	42
2.3.2	Southern Blot Solutions	42
2.3.3	Protein Gel Solutions	43
2.3.4	$\beta$ -galactosidase Staining Solutions	44
2.3.5	Solutions for <i>lacZ</i> Assays	44
2.3.6	Solutions for the Preparation of Ovarian Nuclear Extracts	45
2.3.7	Gel Retardation Assay Solutions	45
2.4	Methods	46
2.4.1	Microbiological Techniques	46
2.4.1.1	Growth of <i>E.coli</i> Bacterial Cultures	46
2.4.1.2	Storage of <i>E.coli</i> Bacterial Cultures	46
2.4.1.3	Preparation and Transformation of Competent Cells	47
	a) $\text{CaCl}_2$ Method	47
	b) DMSO Method	47
	c) Electroporation Method	48
2.4.1.4	Small Scale Preparation of Plasmid DNA	49
	a) Alkaline Lysis Method	49
	b) Rapid Boiling Method	50

2.4.1.5	Large Scale Preparation of Plasmid DNA	50
a)	CsCl <sub>2</sub> Method	50
b)	Qiagen Plasmid Maxi Kit	51
2.4.1.6	Estimation of Plasmid DNA Yields	52
2.4.2	Manipulation and Detection of Nucleic Acids	52
2.4.2.1	Phenol and Chloroform Extraction of Proteins from Nucleic Acid	52
2.4.2.2	Precipitation of DNA	52
2.4.2.3	Restriction Endonuclease Digestion of DNA	53
2.4.2.4	Agarose Gel Electrophoresis of DNA	53
2.4.2.5	Molecular Weight Markers	54
2.4.2.6	Recovery of DNA from Agarose Gels	54
a)	GeneClean	54
b)	Qiaex DNA Gel Extraction	55
2.4.2.7	Ligation of DNA Molecules	55
a)	Vector dephosphorylation	55
b)	Creation of Blunt Ends	55
c)	Addition of Linkers to Insert DNA	56
2.4.2.8	Polymerase Chain Reaction	57
2.4.2.9	Dideoxynucleotide Sequencing of Double Stranded Plasmid DNA	57
2.4.2.10	Polyacrylamide Gel Electrophoresis of DNA Sequencing Reactions	58
2.4.3	Manipulations of <i>Drosophila</i> Stocks	59
2.4.3.1	Establishment of Transgenic Lines by P-Element Transformation	59
2.4.3.2	Preparation of Genomic DNA from Adult Flies	60
2.4.3.3	Southern Blotting	61

2.4.3.4	Radioactive Labelling of DNA by Random Priming	62
2.4.3.5	Hybridisation of Southern Blot Filters	62
2.4.3.6	Histochemical Staining of Transformed <i>Drosophila</i>	63
2.4.3.7	Heat Shock Treatment of <i>Drosophila</i>	63
2.4.3.8	20-hydroxyecdysone Treatment of Transformed <i>Drosophila</i>	63
2.4.3.9	Extraction of Haemolymph Samples from <i>Drosophila</i>	63
2.4.3.10	SDS-Polyacrylamide Gel Electrophoresis of Proteins	64
2.4.3.11	<i>lacZ</i> Assays	64
2.4.4	Gel Retardation Assays	65
2.4.4.1	Preparation of Ovarian Nuclear Extracts	65
2.4.4.2	Estimation of Protein Concentration of <i>Drosophila</i> Ovarian Nuclear Extracts	66
2.4.4.3	Radioactive End-labelling of Probe	66
2.4.4.4	DNA-binding Reactions	67
2.4.4.5	Nondenaturing Polyacrylamide Gel Electrophoresis of Protein-DNA Binding Reactions	67
2.4.4.6	Autoradiography	67
CHAPTER 3 - Results		
- Identification of <i>Cis</i> -acting Sequences Regulating <i>yp3</i> Transcription in <i>Drosophila melanogaster</i>		
3.1	Introduction	68
3.2	Analysis of the <i>yp3</i> Gene to Identify <i>Cis</i> -acting Elements Required for Sex-limited, Tissue-Specific Expression	72
3.2.1	The pCaSpeR-AUG- $\beta$ gal Transformation Vector	72
3.2.2	The pERI Transformation Vector	
3.2.3	Construction and Analysis of the <i>yp3</i> Reporter Constructs	75

3.3	Results and Conclusions	77
3.3.1	Isolation of <i>yp3</i> Upstream Sequences that Govern Tissue-Specific Ovarian and Sex-Limited Fat Body Expression of <i>yp3</i>	78
3.3.2	Sequence Comparisons Between the Identified Fat Body Enhancer Elements of <i>yp1</i> , <i>yp2</i> and <i>yp3</i>	80
3.4	Summary and Discussion	81
CHAPTER 4 - Results		
- Gel Retardation Assays		
4.1	Introduction	85
4.2	Gel Retardation Assays with OE3	86
4.3	Sequence Comparisons Between the Identified Ovarian Enhancer Elements of <i>yp1</i> , <i>yp2</i> and <i>yp3</i>	89
4.4	Summary and Discussion	90
CHAPTER 5 - Results		
- The Effects of 20-hydroxyecdysone on <i>yp3</i> Expression		
5.1	Introduction	92
5.1.1	The Effects of 20-hydroxyecdysone on <i>yp1/yp2</i> Gene Expression	93
5.1.2	The Effects of 20-hydroxyecdysone on <i>yp3</i> Gene Expression	95
5.2	<i>yp3</i> Constructs Tested for Responsiveness to 20-hydroxyecdysone	95
5.3	Results and Conclusions	98
5.3.1	Effects of 20-hydroxyecdysone on <i>lacZ</i> Expression in Males Transformed with Reporter Constructs Containing <i>yp3</i> Upstream Sequences	98
5.3.1.1	Response to 20-hydroxyecdysone by <i>yp3</i> Reporter Constructs in the Abdominal Fat Body Tissue	101



5.3.1.2	Response to 20-hydroxyecdysone by <i>yp3</i> Reporter Constructs in the Thoracic Fat Body Tissue	103
5.3.2	Effects of 20-hydroxyecdysone on <i>lacZ</i> Expression in Males Transformed with Constructs Containing <i>yp3</i> Coding and Downstream Sequences	105
5.3.3	Computer Searches to Locate Ecdysone Response Elements in the 4.3 kb <i>yp3</i> Genomic Sequence	106
5.4	Discussion	107
5.4.1	Re-investigation Into the Effects of 20-hydroxyecdysone on <i>yp1</i> and <i>yp2</i> Expression	110
5.4.2	The Effects of Juvenile Hormone on <i>yp</i> Transcription	112
CHAPTER 6 - Final Discussion		
6.1	Enhancer Regions Required for Tissue-Specific <i>yp3</i> Expression	113
6.2	<i>Trans</i> -acting Factors Mediating <i>yp3</i> Fat Body Transcription	114
6.3	<i>Trans</i> -acting Factors Mediating Ovarian <i>yp3</i> Transcription	114
6.4	The Hormonal Control of <i>yp</i> Transcription	115
6.5	Concluding Remarks	117
REFERENCES		119

## **CHAPTER 1**

### **Introduction**

#### **The Yolk Proteins and Their Genes in *Drosophila melanogaster***

## 1.1 Oogenesis

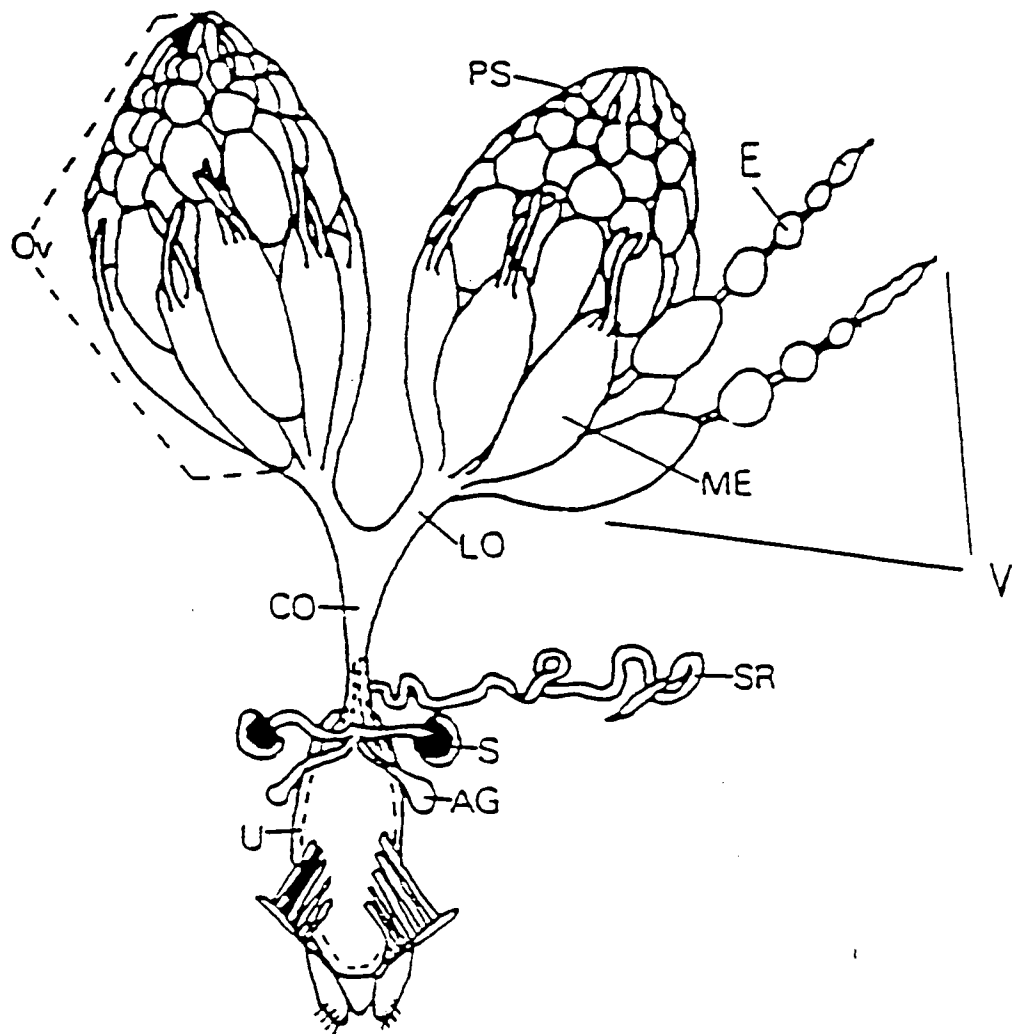
The development of mature oocytes in *Drosophila melanogaster* has been extensively studied and well characterised. Females have one pair of ovaries connected by a common oviduct and each ovary consists of about 15-20 ovarioles lying in parallel (Figure 1.1). At the anterior tip of each ovariole is the germarium where germ line stem cells divide to produce two daughter cells. One of these remains a stem cell and the other goes through four mitoses with incomplete cell division to form a cluster of 16 cells connected by intercellular bridges (Mahowald and Strassheim, 1970). One of these 16 cells is destined to develop into the oocyte while the other 15 will become nurse cells. Before this cluster of cells exit from the germarium into the vitellarium they are surrounded by small somatic profollicle cells, to form the egg chamber. As an egg chamber develops, the volume of the oocyte is greatly increased by the synthesis and uptake of yolk protein components, a process termed vitellogenesis. The morphology of the egg chamber changes dramatically throughout oogenesis and this process has been classified into 14 stages (King, 1970) that depend on the size, shape and appearance of the egg chambers (Figure 1.2). The selective uptake of yolk occurs during stages 8-11 (inclusive) of oogenesis by receptor-mediated endocytosis. This ceases after stage 11, when the vitelline membrane and chorion are assembled and yolk proteins can no longer reach the oocyte surface. The oocyte also increases in size due to the uptake of products from the nurse cells. During oogenesis the nurse cells undergo DNA replication and become polyploid (Jacob and Stirling, 1959) which possibly reflects their function of providing the oocyte with essential components such as lipid droplets (King, 1970) and several RNA species, most of which is ribosomal (Mahowald and Tiefert, 1970) and some of which is mRNA involved in determining egg polarity and subsequent aspects of embryonic development (reviewed by Lasko, 1994).

The follicle cells are the third cell type that compose the egg chamber, which differentiate from the somatic profollicle cells that envelop the nurse cells and oocyte in the germarium (reviewed by Margaritis et al., 1980). At stage 2 of oogenesis there are approximately 80 follicle cell precursors (King, 1970; Bownes

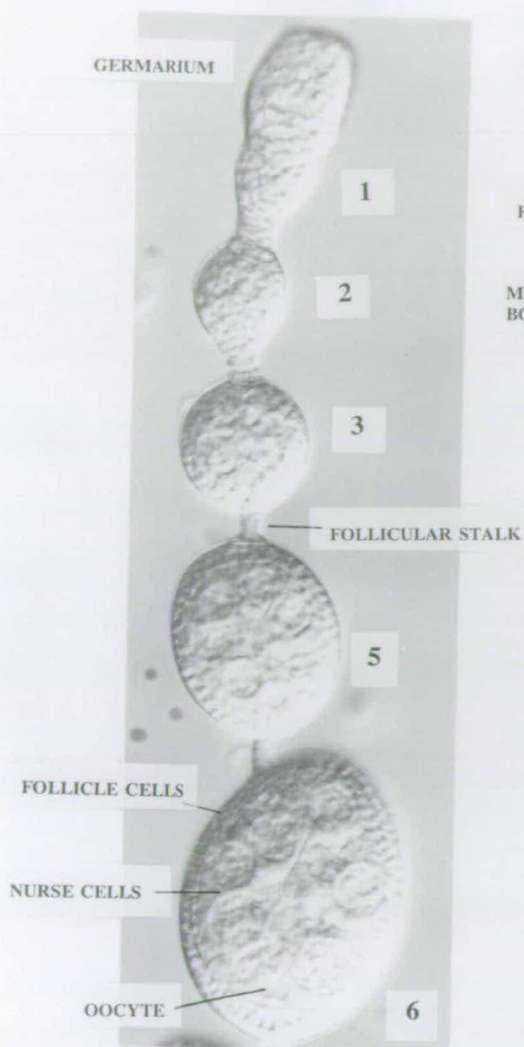
## Figure 1.1

The adult female reproductive system of *Drosophila melanogaster*

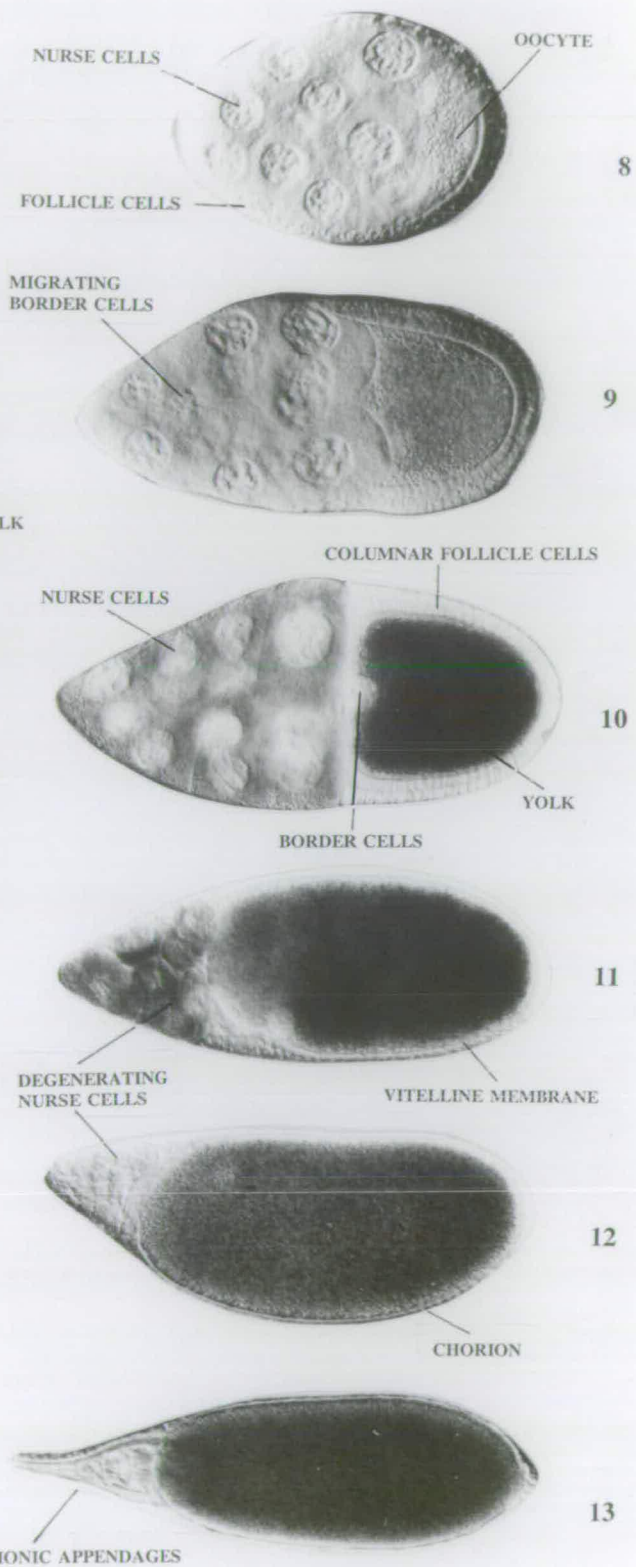
PS	- peritoneal sheath
E	- egg chamber
Ov	- ovary
V	- ovariole
ME	- mature egg
LO	- lateral oviduct
CO	- common oviduct
SR	- seminal receptacle
S	- spermatheca
AG	- accessory gland
U	- uterus



## PREVITELLOGENIC STAGES



## VITELLOGENIC STAGES



and Dale, 1982). Four mitotic divisions occur during stages 2 to 5 to produce approximately 1000 follicular epithelial cells which subsequently differentiate into three follicle cell subtypes, namely the border cells, cells covering the nurse cells and the columnar cells, all of which undergo a series of cell movements during stages 7 to 13 of oogenesis (King, 1970; Logan and Wensink, 1990; Figure 1.3A and B). During stages 7 to 10 of oogenesis the border cells, of which there are 6-10, migrate from the anterior pole of the egg chamber between the nurse cells to reach the anterior surface of the oocyte, and are later involved in the formation of the micropyle through which fertilization occurs (King, 1970). At stages 8 to 9, there are about 50-100 cells surrounding the nurse cells whose function remains undetermined. During stages 8 to 11 the columnar cells surrounding the posterior surface of the oocyte migrate to surround the expanding oocyte and differentiate into anterior pole cells, posterior pole cells and columnar main body cells. The anterior pole cells intercalate between the nurse cells and the oocyte and are thought to be involved in the formation of the respiratory appendages of the mature oocyte, and the posterior pole cells are thought to form the aeropyle which may be involved in embryonic respiration. The columnar main body cells secrete the protective vitelline membrane (stages 8 to 10) (Savant and Waring, 1989; Gigliotti et al., 1989; Fagnoli and Waring, 1982) and chorion (stages 10b to 14) (reviewed in Mahowald and Kambyzellis, 1980) that surround the mature oocyte. At stage 14 of oogenesis, chorion synthesis is complete and these follicle cells degenerate. Importantly, the follicle cells derived from the columnar cells are also involved in yolk protein synthesis, during stages 8 to 10 of oogenesis, which is essential for oocyte maturation. The expression and secretion of yolk proteins, vitelline membrane proteins and chorion proteins by the follicle cells is necessary to complete the formation of the egg. The isolation and molecular analysis of genes encoding these proteins has revealed that this sequence of events involves *cis*-acting regulatory elements required for their correct pattern of expression (discussed in sections 1.5.2 and 1.5.4)

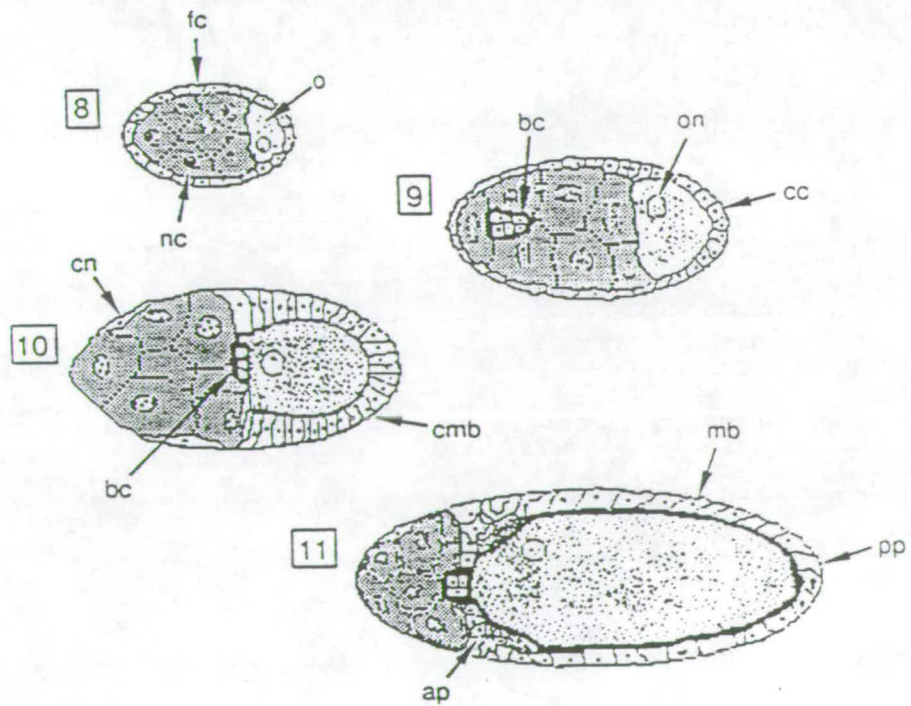
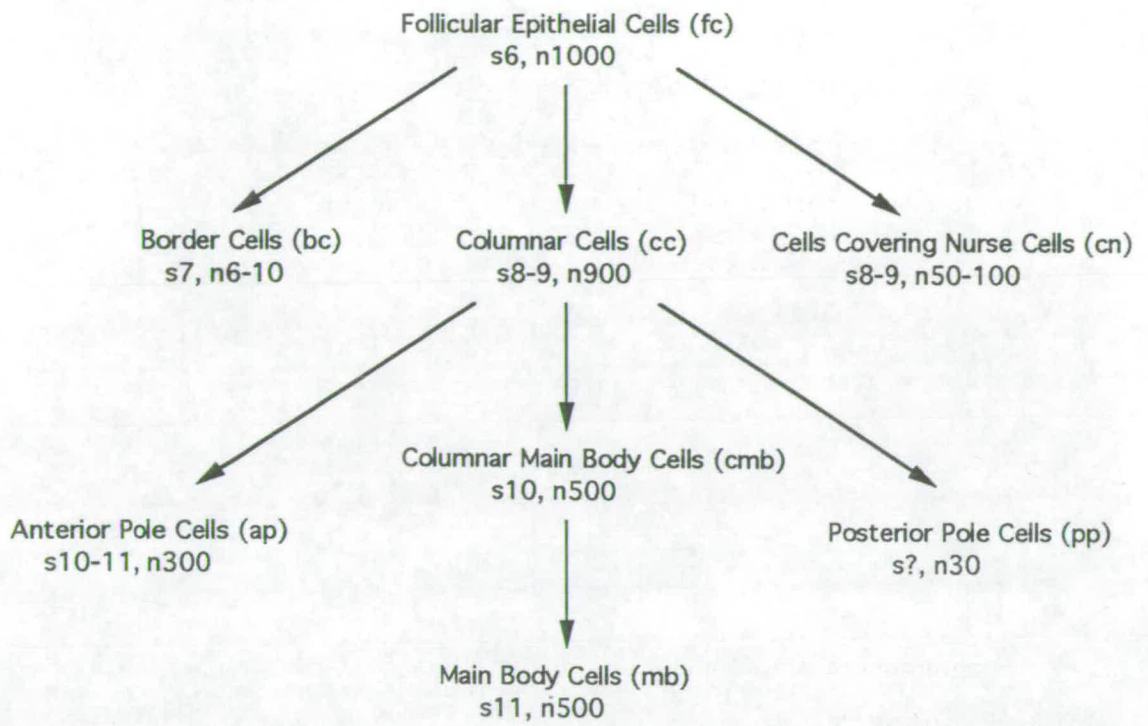
### Figure 1.3

**A** The differentiation of *Drosophila* ovarian follicle cells (Logan and Wensink, 1990). Each differentiation lineage is indicated by arrows. The approximate developmental stage of appearance (s) and approximate number of cells (n) are listed under each cell type name

**B** Diagrammatic representation of sectioned egg chambers (based on King, 1970 and adapted from Logan and Wensink, 1990). The boxed numbers indicate the developmental stage of oogenesis

- o - oocyte
- fc - follicle cells
- nc - nurse cells
- bc - border cells
- on - oocyte nucleus
- cc - columnar cells
- cn - cells covering nurse cells
- cmb - columnar main body cells
- ap - anterior pole cells
- mb - main body cells
- pp - posterior pole cells





## 1.2 Vitellogenesis

Vitellogenesis is the term given to the process of yolk protein synthesis and the accumulation, or specific uptake, of yolk into the developing oocyte. The process, from the beginning of yolk accumulation to egg deposition takes about 20 hours at 25°C (King, 1970). In *Drosophila* there are three major yolk proteins (YP1, YP2 and YP3) which are synthesized in the fat body of the adult female (a tissue often considered to be equivalent to the liver in vertebrates) and in the ovarian follicle cells surrounding the developing oocyte during stages 8 to 10 of oogenesis (Bownes and Hames, 1978; Brennan et al., 1982; Isaac and Bownes, 1982). Yolk proteins synthesized in the fat body are secreted into the haemolymph and selectively endocytosed into the oocyte. Transcripts are first detected soon after eclosion and increase to reach a maximum after 24 hours (Barnett and Wensink, 1981; Isaac and Bownes, 1982). This correlates closely with yolk protein synthesis as the YPs appear in the haemolymph shortly after eclosion, and in the oocyte approximately 10 hours later (Bownes and Dale, 1982). YPs synthesized by the follicle cells are secreted unidirectionally towards the oocyte and become directly associated with the oocyte surface. Regardless of their site of synthesis, the yolk proteins are taken up into the oocyte by receptor-mediated endocytosis and assembled into yolk granules or spheres classified as  $\alpha$ -spheres and  $\beta$ -spheres. The formation of  $\alpha$ -spheres is continuous from and including stages 8 to 10 of oogenesis and they are composed of crystalline protein which appears to be the storage state of the YPs (Butterworth et al., 1991). In comparison, the  $\beta$ -spheres are formed mainly in the later stages 13 to 14 of oogenesis, and contain glycogen (Bownes, 1982). The majority of the yolk in the  $\alpha$ -spheres is composed of YP1, YP2 and YP3 (Bownes and Hames, 1977; Warren and Mahowald, 1979). In their native form, these yolk proteins are known to form oligomers that have been proposed to be hexamers (Isaac.P.G, PhD Thesis, 1982) or tetramers (Fourney et al., 1982). Yolk protein synthesis is integral to vitellogenesis and is thereby essential for oocyte development. The yolk proteins have therefore been investigated directly at the biochemical level, to identify regions important for their uptake into the oocyte, and

extensively at the molecular level in attempts to determine the requirements for *yp* gene expression. Several physiological factors are known to influence vitellogenesis such as nutrition, the insect hormones 20-hydroxyecdysone and juvenile hormone, and the sex-determination genes, which will also be discussed. Experiments to investigate how these factors affect vitellogenesis have revealed that levels of *yp* transcription and protein synthesis are similarly affected, suggesting that regulation occurs by affecting transcript stability and/or transcription rates. There is no evidence to suggest that any translational control of *yp* mRNA operates.

### **1.3 Biochemical Analysis of the Yolk Proteins**

Since the yolk proteins are known to be selectively taken up into the developing oocyte by receptor-mediated endocytosis, biochemical studies have been carried out with an aim to identifying the structural requirements or domains of the YPs that may be involved in receptor recognition. The yolk proteins do not appear to differ in their biochemical properties whether they are synthesized in the ovarian or fat body tissue, although there appear to be differences in their rates of synthesis in these tissues. It was demonstrated that all three yolk proteins were present in equivalent amounts in the fat body cells, but in the ovary YP3 synthesis was reduced to approximately 25 % that of YP1 and YP2 (Brennan et al., 1982; Isaac and Bownes, 1982). SDS-PAGE analysis of the products of cell free translations from fat body and ovarian mRNA showed that the *yp3* transcript was under-represented in the ovary relative to the amount of *yp1* and *yp2* transcripts (Isaac and Bownes, 1982). This result was confirmed by Northern blots using RNA from ovarian and fat body tissue. It was possible that this could result from the selective amplification of *yp1* and *yp2*, since the chorion genes are amplified in the follicle cells prior to their transcription (reviewed by Orr-Weaver, 1991), but no evidence was found to support this theory. However, it was discovered that the rate of transcription was the same for all three *yp* genes, indicating that the observed levels of *yp3* mRNA in the ovary may result from reduced transcript stability (Williams and Bownes, 1986).

All three yolk polypeptides have signal peptides characteristic of secreted proteins which are translationally cleaved prior to secretion from the fat body (Brennan et al., 1980). The signal peptide cleavage sites are thought to be between residues 19 and 20 in all three proteins (Yan et al., 1987; Liddell.S, Ph.D Thesis, 1989). The YPs also undergo post-translational modifications of phosphorylation, glycosylation and sulphation. Phosphorylation can alter the activity of proteins and the YPs are known to be phosphorylated *in vivo* to an equal extent at both sites of synthesis (Brennan and Mahowald, 1982; Minoo and Postlethwait, 1985). However, it is not clear where the sites of phosphate addition are, or at what stage during yolk protein maturation this occurs.

Evidence for *in-vivo* glycosylation of the yolk proteins came from the injection of females with radiolabelled  $^{14}\text{C}$ -mannose or  $^{14}\text{C}$ -glucosamine sugars, which were incorporated into the mature YPs (Minoo and Postlethwait, 1985). There are two modes of glycosylation, N-glycosylation and O-glycosylation according to the amino acid residues at which this takes place. The defined site for N-glycosylation is Asn-X-Ser/Thr, but of the three YPs only YP2 has a potential site at residue 25 where the amino acid sequence is Asn<sup>25</sup>-Arg-Ser (Hung and Wensink, 1983; Garabedian et al., 1987; Yan et al., 1987). If the YPs are glycosylated, it would therefore appear to be through O-glycosylation which takes place at serine or threonine residues and this is further supported from the treatment of radiolabelled YPs with endoglycosidase-H (which cleaves N-linked sugars) after which they remained labelled (Minoo and Postlethwait, 1985). However, with the *in-vivo* labelling of the YPs, it was never shown that the incorporated radiolabel was solely due to the injected radiolabelled monosaccharides. These monosaccharides may have been recycled and the resulting radiolabelled amino acids incorporated into newly synthesized YPs. The evidence for glycosylation of all three YPs is not conclusive since analysis of the carbohydrate content of YPs purified from *Drosophila* eggs showed that carbohydrate was only detectable in YP2 (Mintzas and Kambyzellis, 1982). Glycosylation of proteins is thought to stabilise protein interactions and/or to provide receptor recognition sites. More detailed studies will

be necessary to determine if all the yolk polypeptides are glycosylated and if so, the actual sites and modes of glycosylation.

The sulphation of proteins occurs at tyrosine residues and is a non-reversible highly specific reaction catalysed by the enzyme tyrosylprotein sulphotransferase (Huttner and Baeuerle, 1988). The sulphation of *Drosophila* YPs was demonstrated by the treatment of females with inorganic <sup>35</sup>S-sulphate, after which they were prominently labelled (Baeuerle and Huttner, 1985). The site of tyrosine sulphation in YP2 was later identified as tyrosine 172 (Baeuerle et al., 1988) and removing the sulphation site by changing this residue to phenylalanine 172 inhibited secretion of the YP2 protein (Friederich et al., 1988). Tyrosine sulphation is thought to be a means of protecting exposed tyrosines from extracellular oxidants and may also have secondary functions, one of which could be a marker for protein secretion.

More recently, computer searches between the *Drosophila* YPs and two sequenced *Calliphora* yolk protein genes identified the presence of two putative divalent ion-binding sites in the proteins (Martinez and Bownes, 1994). This was of interest as calcium ions had previously been demonstrated to be involved in receptor-mediated endocytosis of yolk proteins into the oocytes of *Manduca* (Osir et al., 1986), *Locusta* (Rohrkasten and Ferenz, 1986), and *Nauphoeta* (Gochoco et al., 1988). It has since been demonstrated that the *Drosophila* YPs bind calcium ions *in vivo* by running the proteins on SDS-PAGE gels in the presence and absence of calcium and EDTA. A change in the position of the YPs is observed when proteins are electrophoresed in the presence of calcium and EDTA (M.Bownes and D. Mauchline, personal communication). It therefore seems that calcium binding could be a conserved feature between the yolk proteins of many species, and possibly required for their specific uptake into the oocyte.

The yolk proteins have always been assumed to provide a nutritional supply of amino acids during embryonic development as they start to be broken down and utilised during the later stages of embryogenesis when midgut formation is occurring (Bownes, 1982). More recently, sequence comparison between the *Drosophila* YPs and other proteins identified a sequence with significant similarity

to the substrate binding sites of vertebrate lipases such as the human lipoprotein lipase, the human hepatic triacylglyceride lipase and the porcine pancreatic triacylglyceride lipase (Bownes et al., 1988b; Persson et al., 1989). The region of shared homology is concentrated in the putative lipid-binding domain and although the YPs do not have any lipase activity, they have been shown to bind ecdysteroids conjugated to fatty acids, since treatment of purified YPs with esterase and proteases results in the release of ecdysteroid hormone. It was therefore proposed that when the YPs are proteolytically cleaved during embryogenesis the conjugates would be available for metabolism to release ecdysone. This would provide a mechanism for the timely release of hormone to trigger other events in embryogenesis, such as cuticle synthesis and secretion (Bownes, et al., 1988b). The embryos of other insects are known to store ecdysone conjugates (Redfern, 1989), and in *Locusta migratoria* embryos, the degradation of polar conjugates bound to vitellin coincides with secretion of the cuticle (Lagueux et al., 1979). It is therefore possible that the YPs may be functionally more important for events during embryogenesis, rather than simply a source of nutrition. This is further supported by the conservation of yolk proteins between other *Drosophila* species, and between other dipteran species such as *Calliphora*, *Lucilia*, *Protophormia* and *Sarcophaga*, which was demonstrated by the selective uptake of yolk proteins from both closely and distantly related species into *D.melanogaster* and *D.funnebris* ovaries (Martinez and Bownes, 1992). This strongly suggests that the domains involved in receptor-mediated endocytosis are conserved, and that other features such as the putative domain for binding ecdysteroid conjugates could also be conserved.

## **1.4 Factors Affecting YP Synthesis and Vitellogenesis**

### **1.4.1 Nutrition**

In *D. melanogaster* vitellogenesis is affected by the nutritional status of the fly. Females that have been starved or fed on a sugar diet lacking other dietary components have reduced levels of vitellogenesis. This is reflected by a decrease in

the rate of egg production which eventually ceases. It is thought that yolk protein uptake and therefore oocyte maturation is affected since egg laying ceases within one day despite the fact that there are still high levels of YP circulating in the haemolymph. Also, it takes approximately two days for *yp* transcript levels to decrease in the fat body. This is not a general effect on transcription rates since levels of  $\alpha$ -tubulin message remain unchanged (Bownes and Blair, 1986), and it does not appear to be the result of under-development of the fat body (Bownes et al., 1988a). This effect is reversible since egg production can be fully restored by returning females to a complete diet.

A similar effect has been reported for the mosquito, *Aedes aegypti*, where vitellogenesis is induced in response to a blood meal (van Handel and Lea, 1984). These observations suggest that a system exists to modulate *yp* expression and egg production in relation to nutrition.

## **1.4.2 Hormones**

### **1.4.2.1 Juvenile Hormone**

For several years the insect hormones juvenile hormone (a sesquiterperoid hormone) and ecdysone (a steroid hormone) have been known to influence *yolk protein* gene expression. Early experiments on the isolated abdomens of newly eclosed females, which have reduced levels of YP synthesis, demonstrated that treatment with ZR515 (a synthetic analogue of juvenile hormone) or 20-hydroxyecdysone (the physiologically active form of ecdysone) stimulated fat body synthesis and secretion of YPs (Jowett and Postlethwait, 1980; Wu et al., 1987).

Other experiments with juvenile hormone (JH) have shown that levels of *yp* transcription are increased in the ovary and fat body of starved adult females which have reduced levels of *yp* expression. There is a direct relationship between nutritional intake and levels of *yp* transcription, which subsequently affects the rate of egg production in females (Bownes and Blair, 1986; Bownes et al., 1988a). In the ovary, JH stimulates the uptake of yolk proteins into the oocyte and

thus the passage of oocytes through vitellogenesis (Postlethwait and Handler, 1978; Jowett and Postlethwait, 1980; Bownes, 1986) and it appears that increased *yp* expression is due to egg chambers entering the correct developmental stages for *yp* transcription to occur, rather than a direct effect of JH on the *yp* genes themselves. More recently, the ultrastructure of the oocyte membrane has been shown to be altered by ZR515 *in vitro* such that the oocyte becomes competent to endocytose material from the culture medium, supporting the hypothesis that JH increases *yp* synthesis in the ovary by encouraging the uptake of YPs circulating in the haemolymph and thus oocyte maturation (Giorgi et al., 1993). *In vivo* evidence is provided by studies of *apterous*<sup>4</sup> (*ap*<sup>4</sup>) mutants which have low levels of JH due to a defective corpora allata and hence fail to undergo ovarian development and vitellogenesis (Bownes, 1989). Nevertheless, there is still some YP synthesis and many of these flies have normal levels of YP circulating in their haemolymph, but endocytosis into the oocyte is only stimulated by treatment with ZR515 (Postlethwait and Weiser, 1973; Redfern and Bownes, 1982).

It is not clear how ZR515 increases *yp* transcript levels in the fat body of starved or normally fed females (Bownes et al., 1987). Attempts to identify *cis*-acting *yp* sequences necessary for these effects have all been negative. Several constructs containing fragments of *yp1*, *yp2* or *yp3* fused to *Adh* or *lacZ* reporter genes did not show increased expression following treatment with ZR515. An increase in *yp* transcript levels is only observed with the native *yp* genes suggesting that JH response elements, if they exist, are located within *yp* coding sequences or introns (Bownes et al., submitted). These results are compatible with a recent report on the isolation of a putative JH receptor (JHR) from larvae of the tobacco hornworm (*Manduca sexta*) indicating that the effects of JH on gene expression are indirect. The receptor is a 29 kDa nuclear protein that specifically binds JH with high affinity and has a developmental pattern of expression that coincides with, and is required for periods of JH action. The JHR is present during larval growth periods and both JH and JHR are required at the time of ecdysteroid induced molting, such that the response to ecdysteroid results in a molt, rather than metamorphosis. At the time of pupal commitment and in the absence of JH, JHR disappears in response to



ecdysteroid. It reappears in pupae when the cells again become responsive to JH. This information provides strong evidence that the protein is a JH receptor, yet it remains unclear how the JH-JHR complex mediates its effects. Although the receptor is localized to the nucleus, sequence analysis has shown that it has no known DNA-binding motifs and it cannot be classified as a member of any known family of hormone receptors (Palli et al., 1994). A molecular mechanism for the action of JH therefore remains undetermined, although the identification and isolation of a JHR has indicated that it is unlikely to be at the level of DNA-binding. Further studies should reveal the function of JHR, whether it is present in other insect species, and if so the degree of conservation which would subsequently help to identify regions of importance for JH-JHR activity.

An equivalent to the *Manduca* JHR is yet to be identified in *Drosophila* and in the adult, where JH is involved in regulating various aspects of reproductive maturation, including vitellogenesis and thus levels of *yolk protein* gene expression (Bownes, 1986; Wyatt, 1991). Some time ago, a search was carried out for genes involved in the transactivation of *yp* expression by screening for mutants that failed to drive expression from a *yp-Adh* gene fusion. In this screen, the mutant *cricketlet* (*clt*) was identified that was defective in *yp* synthesis and showed defective histolysis of the larval fat body and had rudimentary ovarian development. This phenotype suggested that *clt* could be defective somewhere in the JH response pathway and may have been mutant for a JH carrier protein or receptor (Shirras and Bownes, 1989).

Since only the native *yp* genes are affected by JH it is reasonable to postulate that juvenile hormone may enhance transcript levels by improving transcript stability. If the *Manduca* JHR is a conserved protein, it is possible that a JH-JHR complex could influence RNA metabolism in *Drosophila* by affecting post-transcriptional RNA processing or RNA translation, since the putative *Manduca* JHR has some similarity to an RNA helicase and is present in nuclear structures associated with the nucleolus (Palli et al., 1994).

#### 1.4.2.2 20-Hydroxyecdysone

It seems that the molecular action of ecdysone will be more straightforward to understand in *Drosophila* since it appears to act upon target genes directly. Preliminary evidence for this came from studies on the steroid induced puffing of larval salivary gland polytene chromosomes, where ecdysteroid initiates a cascade of gene activation and inactivation. This led to the Ashburner model for ecdysteroid action which involved a putative ecdysteroid receptor such that an ecdysteroid-receptor complex could differentially regulate two classes of target genes, termed the "early" and "late" genes. It was proposed that the early genes were directly activated by the ecdysteroid-receptor complex whilst the late genes were inactivated, and that the early genes encoded regulatory factors that could induce the observed secondary response to the hormone, where the late genes were activated and the early genes repressed (Ashburner et al., 1974; Ashburner and Richards, 1976). This theory is supported by a recent molecular analysis of three early ecdysteroid responsive genes in *Drosophila*, namely E74 (Burtis et al., 1990; Thummel et al., 1990), E75 (Segraves and Hogness, 1990) and the *Broad* complex (DiBello et al., 1991) which encode multiple DNA binding proteins with properties of transcription factors (reviewed by Andres and Thummel, 1992). The different forms arise from the use of alternative promoters and/or alternative splicing and their expression does not appear to be developmentally regulated or tissue-specific. Instead, the appearance of different forms appears to depend upon the sensitivity of the alternative promoters to 20-hydroxyecdysone and transcript length (Karim and Thummel, 1991, 1992). The early genes are longer than most other known *Drosophila* genes (60-100 kb) such that there could be a significant delay in the synthesis of primary transcripts which would affect the overall timing of gene expression in a regulatory hierarchy.

A gene encoding the *Drosophila* ecdysone receptor (*EcR*) has also been recently isolated and characterized, and the *EcR* protein classified as a member of the steroid receptor superfamily (Koelle et al., 1991). Members of this family share a highly conserved DNA-binding domain for the recognition of specific DNA

sequences, or hormone response elements. The carboxy-terminal region is less conserved, containing sequences for ligand binding and dimerization (Evans, 1988; Forman and Samuels, 1990; Laudet et al., 1992). A putative consensus sequence for the ecdysone response element in *Drosophila* was proposed following studies on the ecdysone inducible *Eip28/29* and *hsp27* genes. The *Eip28/29* gene was isolated in a biochemical screen for genes that responded rapidly when Kc cells (an embryonic cell line derived from first instar larvae) were treated with 20-hydroxyecdysone (Savakis et al., 1980, 1984). The *hsp27* gene was also investigated since it was known that prepupal expression could be induced by 20-hydroxyecdysone as well as heat shock (Ireland and Berger, 1982; Morganelli et al., 1985). A detailed promoter analysis of these genes isolated the DNA elements required for ecdysone induction and the consensus ecdysone response element was found to be an imperfect palindrome with the sequence RGG/TTCA<sub>n</sub>TGAC/ACY, where R represents G or A, Y represents T or C and n any nucleotide (Riddihough and Pelham, 1987, Cherbas et al., 1991; Martinez et al., 1991). This is consistent with the observation that steroid receptors usually act as dimers. The ecdysone receptor appears to act only as a heterodimer with the *ultraspiracle* protein (USP) which also belongs to the steroid receptor superfamily (Yao et al., 1992; Thomas et al., 1993). The USP protein was initially identified as Chorion Factor 1 (CF1) from binding to the promoter region of the *Drosophila* *s15* chorion gene (Shea et al., 1990), and was later identified to be USP, a homologue of the vertebrate retinoid X receptor (Christianson et al., 1992; Christianson and Kafatos, 1993). These receptors bind to hormone response elements to modulate the transcription of specific target genes.

More recently, it was reported that there are at least three different functional isoforms of the *Drosophila* ecdysone receptor (EcR-A, EcR-B1 and EcR-B2) that share common DNA- and hormone-binding domains and are distinguished by their different N-terminal regions thought to be important for the transactivation of target genes (Talbot et al., 1993). The information obtained from these studies has made it possible to understand, to some extent, how 20-hydroxyecdysone can have such diverse effects on different tissues during *Drosophila* development. During the first and second larval instars, a pulse of 20-hydroxyecdysone triggers molting of the

molting of the larval cuticle, but at the end of the third instar the response is entirely different. A pulse of 20-hydroxyecdysone signals puparium formation and the onset of metamorphosis, during which the imaginal tissues of the larva are stimulated to start differentiating into adult structures, whilst the strictly larval tissues that provided the functional structures of the larva undergo histolysis and degenerate. The first hypothesis to explain these different responses to 20-hydroxyecdysone was the tissue coordination model. This was basically an extension of the Ashburner model since it was discovered that the early gene transcription factors could also be ecdysone induced in larval and imaginal tissues, demonstrating that a response to this hormone was not limited to the larval salivary gland. It was proposed that different target tissues had different metamorphic responses to ecdysone due to a cascade of gene activation and inactivation, which could result from different combinations of early gene transcription factors produced by a similar genetic regulatory hierarchy to that discovered in the larval salivary gland (Burtis et al., 1990; Thummel et al., 1990). With the more recent isolation of the ecdysone receptor protein and the identification of the various isoforms, this model was extended further. It was proposed that different stage- and tissue-specific responses to 20-hydroxyecdysone could be mediated through the use of different stage- and tissue-specific ecdysone receptors. This was supported by experiments with isoform-specific monoclonal antibodies which revealed that different ecdysone target tissues did express different combinations of receptor isoforms during metamorphosis (Talbot et al., 1993). Thus, instead of various combinations of early gene transcription factors regulating the fate of larval and imaginal tissues at metamorphosis, it is thought that each of the receptor isoforms interacts with 20-hydroxyecdysone to affect particular genetic regulatory hierarchies that will result in a specific metamorphic response. Protein products of such genetic hierarchies could interact with tissue-specific factors or even with the ecdysteroid-receptor complex itself to regulate the expression of structural genes. With such a mechanism through which to operate, it can be envisaged how vastly different responses to 20-hydroxyecdysone can occur during different developmental stages and within different tissues present at these stages.

Much less is known about the effects of 20-hydroxyecdysone in the adult, although it is known to be involved in regulating adult reproduction. As part of this system, vitellogenesis may be one of the processes affected which could be regulated by controlling levels of *yp* gene expression. The effects of 20-hydroxyecdysone on *yp* expression have therefore been extensively investigated. In comparison to JH, 20-hydroxyecdysone only stimulates expression of *yps* in the fat body and there is no evidence to suggest that *yp* transcripts are altered in the ovary or any other adult tissues (Bownes et al., 1983). As previously mentioned, initial experiments on isolated abdomens demonstrated that treatment with 20-hydroxyecdysone increased YP synthesis (Jowett and Postlethwait, 1980; Wu et al., 1987). A similar effect is also observed in the fat body tissue of starved females treated with this hormone (Bownes and Blair, 1986). The most noticeable induction of *yp* expression occurs in males. Following injection with 20-hydroxyecdysone YPs can be detected in the haemolymph (Postlethwait et al., 1980) and *yp* transcripts appear in the fat body (Bownes et al., 1983). This response is transient (Bownes, 1982) and appears to override the sex-determination hierarchy which prevents female differentiation genes, of which the *yolk protein* genes are an example, from being expressed in males. Pseudomales or intersexual flies which have mutations in either *tra* or *dsx*, but do not normally synthesize YPs, can be induced to do so after the application of 20-hydroxyecdysone (Bownes and Nöthiger, 1981; Bownes et al., 1987). It has been shown that the induction of *yp* transcription in males by 20-hydroxyecdysone can be prevented by cycloheximide, an inhibitor of protein synthesis, suggesting that at least one other gene product is required for this activation (Bownes et al., 1987), one of which is likely to be the ecdysone receptor.

Although *yp* gene expression is affected by 20-hydroxyecdysone under these experimental conditions, the *in vivo* role of ecdysone in the regulation of YP synthesis remains unclear. Techniques have not been available to accurately measure levels of ecdysterone in adult *Drosophila* and there have been reports that levels are similar in males and females and do not alter throughout the adult lifespan (Bownes and Rembold, 1987; Bownes, 1989, 1990). It has also been reported that variations in ecdysone levels do occur (Schwartz et al., 1985). However, in other

diptera such as *Calliphora* and *Musca*, a protein meal raises levels of ecdysone and triggers cyclic vitellogenesis, and in the mosquito, *Aedes aegypti*, titres of 20-hydroxyecdysone increase after a blood meal (Racciopi et al., 1986) and vitellogenesis is induced (van Handel and Lea, 1984).

In adult *Drosophila*, the levels of hormone required to induce YP synthesis in males or increase levels in females are several orders of magnitude greater than the normal hormone titre. However, the fact that *yp* expression can be stimulated by 20-hydroxyecdysone suggests that these genes contain ecdysone response elements, and the ability to induce *yp* expression in males provides an ideal system in which to identify the presence of such elements in *yp* constructs fused to reporter genes. Such an analysis has been carried out on all three *yp* genes and is discussed in detail in Chapter 5.

### 1.4.3 Sex-determination Genes

#### 1.4.3.1 Genetic Studies

A primary factor determining YP synthesis in *Drosophila melanogaster* is the sex of the fly. The initial determinant of sexual phenotype is the ratio of X chromosomes compared to autosomes in a cell. A ratio of 1 (XX:AA) results in female differentiation, and a ratio of 0.5 (XY:AA) results in male differentiation. The YP's are synthesised in the fat body of adult females but not males. From genetic studies it is known that sexual dimorphism in *Drosophila* is governed by a hierarchy of somatic sex-determination genes (Figure 1.4) that include *Sex-lethal* (*Sxl*), *transformer* (*tra*), *transformer-2* (*tra-2*), *doublesex* (*dsx*) and *intersex* (*ix*) (reviewed by Wolfner, 1988; Baker, 1989; Belote, 1989; Slee and Bownes, 1990; Steinmann-Zwicky et al., 1990). *Sxl* is active only in females and is dependent on the products of two maternally acting genes, *daughterless* (*da*) (Cline, 1978) and *Daughter killer* (*Dk*). Both these genes have lethal temperature sensitive mutations and they are thought to act together to activate *Sxl* in female embryos. *Sxl* activation determines a commitment to the female differentiation pathway and it is

#### Figure 1.4

The somatic sex determination pathway in *Drosophila melanogaster*. A number of zygotic and maternally supplied products initiate the correct splice form of *Sxl* (see text for details). Female-specific splicing of the *Sxl* transcript results in an active protein that dictates the female-specific splicing pattern of *tra* which in turn dictates the female-specific splicing pattern of *dsx*. Male-specific splicing of the *Sxl* transcript results in a truncated reading frame and the production of an inactive protein. In the absence of an active *Sxl* protein product, default splicing of the *tra* transcript occurs, again resulting in a truncated reading frame and the production of an inactive protein. The production of an inactive *tra* protein in turn dictates the male-specific splicing pattern of *dsx*. The male- and female-specific *dsx* transcripts both encode active sex-specific proteins which subsequently influence sexual differentiation

X CHROMOSOME  
NUMERATORS

*sis-a*  
*sis-b*  
*sis-c*  
*runt*

MATERNALLY  
SUPPLIED

*da*  
*Dk*  
*snf*

PROMOTE FEMALE-  
SPECIFIC *Sxl*  
SPLICING

XX:AA

MATERNALLY  
SUPPLIED

*emc*

AUTOSOMAL  
DENOMINATORS

*deadpan*

REPRESS FEMALE-  
SPECIFIC *Sxl*  
SPLICING

XY:AA

*Sxl* ACTIVE

*Sxl*

*Sxl* INACTIVE

*tra* ACTIVE

*tra*

*tra* INACTIVE

*dsx* ACTIVE

*dsx*

*dsx* ACTIVE

ACTIVE PROTEIN  
PRODUCT

DSX<sup>F</sup>

ACTIVE PROTEIN  
PRODUCT

DSX<sup>M</sup>

MALE DIFFERENTIATION  
GENES REPRESSED

FEMALE DIFFERENTIATION  
GENES REPRESSED



also required for dosage compensation of X-linked genes. In females it acts as a negative regulator of dosage compensation genes which hyperactivate X-linked gene transcription in males (Cline, 1983, 1984; Gergen, 1987). A gene identified as a possible mediator of the functions of the *Sxl* gene in dosage compensation and sex determination is *virilizer* (*vir*) as temperature sensitive mutant females develop as intersexes (Hilfiker and Nöthiger, 1991). *vir* is known to act upstream of *tra* since constitutive expression of *tra* in *vir* mutants leads to the development of sterile females, therefore rescuing the mutant phenotype.

Analysis of *Drosophila Sxl*, *tra*, *tra-2* and *dsx* mutants has demonstrated that the genes in the sex-determination hierarchy directly influence *yp* gene expression and subsequent YP synthesis (Postlethwaite et al., 1980; Bownes and Nöthiger, 1981). XX chromosomally female flies, but mutant at either *tra* or *tra-2* (e.g. X/X ; *tra/tra*) do not express the *yp* genes and develop morphologically as males, termed pseudomales. XX chromosomally female flies mutant at the *dsx* locus (X/X ; *dsx/dsx*) develop as intersexes displaying both male and female phenotypes and, as such, exhibit varying levels of YP synthesis correlating to the degree of male or female characteristics exhibited. Flies with a predominantly male phenotype have the lowest levels of YP synthesis. XY chromosomally male flies mutant at these loci (e.g. X/Y ; *dsx/dsx*) develop as intersexes and are able to synthesize YPs in the fat body.

The effects of the sex-determination hierarchy on *yp* gene expression were further enhanced by experiments with temperature sensitive *tra-2<sup>ts</sup>* mutants (Belote et al., 1985). X/X ; *tra-2<sup>ts</sup>* homozygotes develop as females and undergo YP synthesis at the permissive temperature of 16°C, but develop as pseudomales which do not undergo YP synthesis when reared at the restrictive temperature of 29°C. Further, when X/X ; *tra-2<sup>ts</sup>* females reared at 16°C are shifted to 29°C, *yp* transcription diminishes but can be restored by returning the flies to 16°C (Bownes et al., 1987). This demonstrates quite clearly that the correct expression of the *yp* genes in the fat body depends upon the continuous and correct expression of the genes that constitute the somatic sex determination hierarchy.

This was further supported by the outcome of experiments with a 929 bp *yp1* fragment spanning nucleotides -886 to +43 fused to the *Drosophila Adh* gene. The expression pattern of this construct in transformed flies was determined by measuring Adh activity, which was detected only in the fat body of adult females. When this fusion construct was transformed into flies mutant for *dsx*, no Adh activity was detected in intersexual flies, thereby demonstrating that this *yp* construct contained sites where the *dsx* gene product exerted its effect (Shirras and Bownes, 1987).

*tra-2<sup>ts</sup>* mutants were also used in experiments demonstrating that a different system of *yp* regulation occurs in the ovary. When flies were reared at the permissive female-determining temperature and then shifted to the restrictive temperature, YP synthesis continued in the somatic ovarian follicle cells (Bownes et al., 1987). The sex-determination genes are required to direct the formation of sex-specific tissues at an early developmental stage and once formed, subsequent gene expression appears to be governed by tissue-specific *trans*-acting factors and does not require the continuous action of the sex-determination genes themselves (Wolfner, 1988; Belote, 1992).

#### 1.4.3.2 Molecular Studies

*Sxl* is the gene at the head of the sex-determination pathway and functional transcripts are only produced in females by means of differential splicing of the primary transcript. This results from an assessment of the X:A ratio during embryonic development after blastoderm formation (Baker and Belote, 1983) by several discrete numerator elements for which some of the genes have been identified, namely *sisterless-a* (*sis-a*) (Erickson and Cline, 1993), *sisterless-b* (*sis-b*) which has been identified to be the *scute* gene of the *achaete-scute* complex (Torres and Sánchez, 1989), *sisterless-c* (*sis-c*) (Cline, 1993) and *runt* (Torres and Sánchez, 1992). These are positively acting numerator genes located on the X-chromosome. Other factors involved in assessing the X:A ratio are negatively acting denominators located on the autosomal chromosomes, an example of which is *deadpan* (*dpn*)

(Younger-Shepherd et al., 1992), and maternally supplied products that may activate or suppress the female specific splicing pattern of *Sxl*. Examples of maternally supplied products present in the oocyte known to be positive regulators of female specific *Sxl* splicing are the protein products of *daughterless (da)* (Cronmiller et al., 1988; Caudy et al., 1988; Cabrera and Alonzo, 1991) and *sans-fille (snf)* (Oliver et al., 1988; Salz, 1992), also known as *liz* (Steinmann-Zwicky, 1988). A recent report has identified *snf* to be a putative U1 snRNP (small nuclear riboprotein), a member of the family of proteins known to be involved in forming complexes that recognise the 5' splice sites at intron boundaries. *snf* could therefore be directly involved in the female specific splicing of *Sxl* (Flickinger and Salz, 1994). A negative regulator of *Sxl* is encoded by the *extramacrochaetae (emc)* gene (Van Doren et al., 1991). It is believed that the overall intracellular concentration of these numerator, denominator and maternally supplied factors determines the initial form of *Sxl* splicing since females will have a greater number of X-linked numerator genes. Several of the genes mentioned encode proteins containing helix-loop-helix motifs that are structurally capable of forming heterodimers. It is therefore entirely possible that various heterodimers with differing RNA binding properties could act to direct male- or female-specific splicing of the *Sxl* pre-mRNA (Parkhurst et al., 1990; Younger-Shepherd et al., 1992; Erickson and Cline, 1993).

Ultimately, the X:A ratio therefore determines the functional state of *Sxl* and once set, the female pattern of *Sxl* splicing is maintained by positive autoregulation. *Sxl* has two alternative transcriptional start sites (Keyes et al., 1992) and Northern analysis has revealed that several transcripts are produced from the two promoters. Three transcripts of 4.0 kb, 3.1 kb and 1.7 kb are detected in the early embryo during blastoderm formation that are transcribed from the early promoter. The early *Sxl* protein interacts with other factors to direct a male- or female-specific pattern of splicing from the late promoter. The transcripts produced at this stage are 4.3 kb, 4.2 kb and 1.9 kb. The female-specific *Sxl* product is predicted to be a 354 amino acid protein and the male-specific protein 48 amino acids (Bell et al., 1988).

In females, the resulting active protein contains an RNA binding domain which is involved in directing the female-specific processing of its own

transcript and that of the *tra* pre-mRNA (Bell et al., 1988, 1991; Keyes et al., 1992). This mechanism of RNA processing to produce functional gene products in females is transmitted along the pathway such that the products of *tra* and *tra-2* can positively regulate female specific splicing of *dsx*, the last gene in the hierarchy. The *tra-2* protein product also contains a putative RNA binding domain (Bandziulis et al., 1989; Goralski et al., 1989; Query et al., 1989). The *tra* and *tra-2* proteins both appear to be involved in binding to six repeats of tandemly interspersed 13 bp sequences in the fourth female-specific exon of the *dsx* pre-mRNA (Burtis and Baker, 1989) to direct a pattern of female specific splicing and polyadenylation, since deletion of the region containing these sequences results in a loss of female specific splicing of the *dsx* primary transcript (Hedley and Maniatis, 1991; Hoshijima et al., 1991; Ryner and Baker, 1991). Although the *tra* protein does not appear to have an RNA binding domain, both the gene products of *tra* and *tra-2* do have arginine/serine rich domains which are characteristic of proteins involved in RNA processing (Query et al., 1989) and it has been reported that both the *tra* and *tra-2* proteins bind this *cis* regulatory region *in vitro* (Inoue et al., 1992). The number of 13 bp sequences relates to the efficiency of female specific splicing and evidence from immunoprecipitation studies also indicates that there could be a direct interaction between the *tra* and *tra-2* protein products. However these data need to be confirmed by direct evidence from *in vivo* studies. It also remains to be determined if other sequences in the *dsx* primary transcript are involved in regulating its pattern of alternative splicing. (Hedley and Maniatis, 1991; Hoshijima et al., 1991; Inoue et al., 1992).

In males, most of the sex-determination gene products are non-functional. Male specific splicing of *Sxl* introduces a premature stop codon in the open reading frame resulting in the production of a truncated and inactive protein. This pattern of default splicing is transmitted along the pathway causing male specific splicing of the *tra* transcript which also results in the production of a non-functional protein. A functional *tra-2* mRNA is expressed, however, and required for the correct differentiation of cells in the male germ line (Belote and Baker, 1982, 1983). At the *dsx* locus, a male specific pattern of splicing

occurs that differs from the splicing pattern observed in females, resulting in the production of male-specific and female-specific polypeptides (Burtis and Baker, 1989). Both the male and female DSX proteins (DSX<sup>M</sup> and DSX<sup>F</sup>) are functionally active having different, almost opposite, effects on subsequent sexual differentiation. DSX<sup>F</sup>, in conjunction with the *ix* gene product represses male terminal differentiation genes such that female development occurs. Conversely, DSX<sup>M</sup> represses female differentiation by repressing target female differentiation genes and male development ensues.

#### **1.4.3.3 *doublesex* and the *yolk protein* genes**

The *yolk protein* genes of *Drosophila melanogaster* are the only somatic sexual differentiation target genes known to be under the direct control of the sex-determination hierarchy. As demonstrated by genetic studies (1.4.3.1), their correct expression in the adult fat body requires the continuous action of the genes in this pathway (Postlethwait et al., 1980; Bownes and Nöthiger, 1981; Belote et al., 1985; Bownes et al., 1987; Wolfner, 1988). Ultimately, it appears that the appropriate DSX<sup>M</sup> and DSX<sup>F</sup> proteins must be present in the appropriate sex. Recent molecular studies with *yp1* and *yp2* have demonstrated that there is a direct interaction between the DSX proteins and a tissue-specific enhancer element required for the sex-specific fat body expression of these genes (Burtis et al., 1991). The fat body enhancer (FBE) is a 125 bp sequence located 196 bp upstream of the *yp1* transcription start site and it is required for the sex-specific fat body expression of both *yp1* and *yp2* (Garabedian et al., 1986). This is discussed in more detail in section 1.5.2 that deals with the molecular analysis of the *yolk protein* genes. *In vitro* footprinting studies with the male- and female-specific DSX proteins, expressed in *E.coli*, revealed four binding sites to be present in the FBE (Fig. 1.5). The proposed 9 bp consensus recognition sequence was CTACAAAGT and all protected sites had a homology to this of at least 7 bp (Burtis et al., 1991). It was surprising that both the DSX<sup>M</sup> and DSX<sup>F</sup> proteins bound to the same sequences in the FBE since they have opposite effects on *yp* gene expression in this tissue *in vivo*.

### Figure 1.5

The four protected footprint sequences from binding of DSX proteins to the *yp1/yp2* FBE (adapted from Burtis et al., 1991). Homologies to the proposed consensus binding sequence are underlined, with the first and last nucleotide positions indicated. The number of matched nucleotides to the consensus in each footprint are also given

FOOTPRINT	PROTECTED SEQUENCE	MATCH
A	ACA <u>ACTACAATGTT</u> GCAAT     -303 -295	8/9
B	GAGC <u>CTACAAAGTGATTACAAATT</u> AAAATA         -279 -271 -268 -260	9/9 7/9
C	GGTG <u>CTGCTAAGT</u> CATCA     -235 -227	7/9

Genetic studies had demonstrated that the *yps* were transcribed in *dsx* null mutants, suggesting that DSX<sup>M</sup> prevented *yp* expression in males and DSX<sup>F</sup> had no role in regulating *yp* expression in females (Postlethwait et al., 1980; Bownes and Nöthiger, 1981; Belote et al., 1985). However, sequence analysis of DSX<sup>M</sup> and DSX<sup>F</sup> revealed that these proteins were identical for the first 397 N-terminal amino acids (Burtis and Baker, 1989), a region of which was recently discovered to encode a novel zinc finger DNA binding domain likely to mediate the specific binding to the *yp1/yp2* FBE observed *in vitro* (Erdman and Burtis, 1993). A recent mutational analysis of three of the DSX binding sites in the FBE indicated that DSX<sup>M</sup> repressed and DSX<sup>F</sup> activated *yp* expression from the two strongest binding sites and if both proteins were present (e.g. XX flies with a *dsx<sup>m</sup>/+* genotype) they competed for the same binding sites (Coshigano and Wensink, 1993). DSX<sup>F</sup> could therefore be involved in the activation of *yp* expression in females rather than acting only to repress the expression of male differentiation genes (Jursnich and Burtis, 1993). The different effects of the male- and female-specific DSX proteins on *yp* expression *in vivo* are thought to be mediated by the unique carboxy termini of 152 amino acids in DSX<sup>M</sup> and 30 amino acids in DSX<sup>F</sup>. The mechanisms involved in the regulation of target gene expression are yet to be elucidated and will no doubt prove to be quite complex, involving interactions with tissue specific *trans*-acting factors. The *ix* gene product is also expected to be involved as it is required in conjunction with DSX<sup>F</sup> to regulate target gene expression in females.

It has been demonstrated *in vivo* that the FBE used in the DSX binding studies is not the only region able to confer sex- and fat body-specific expression on the *yp* genes (Abrahamsen et al, 1993) and computer searches reveal putative DSX binding sites in other DNA regions both flanking and constituting the *yp* genes. However the 9 bp sequence occurs frequently throughout *yp1* and *yp2* and the flanking DNA, and also in genes not differentially regulated in males and females such as the alcohol dehydrogenase (*Adh*) and bicoid (*bcd*) genes (Abrahamsen et al., 1993). It is therefore important that putative *in vitro* binding sites are tested by *in vivo* mutational analyses to confirm their authenticity. The identification of additional target genes regulated directly by the DSX proteins



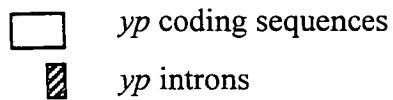
would aid progress towards understanding how DSX<sup>M</sup> and DSX<sup>F</sup> mediate their differing effects on gene expression. A characterisation of other structural domains besides the DNA binding domain would also provide further information about DSX function and could lead to the identification of other regulatory sequences and proteins involved in regulating the expression of specific target genes. To date, there is no evidence to rule out the possibility that either or both of the DSX proteins could interact with other proteins *in vivo* which may alter their binding specificities and subsequently change the recognition of target genes to be regulated.

### 1.5 Molecular Analysis of the *yolk protein* Genes

The yolk proteins of *Drosophila melanogaster* have been extensively investigated at the molecular level. They comprise a small multigene family encoded by three single copy genes on the X chromosome (Postlethwait and Jowett, 1980; Barnett et al., 1980), all of which have been cloned (Barnett et al., 1980) and sequenced (Hovemann et al., 1981; Hung and Wensink, 1981, 1983; Garabedian et al., 1987; Yan et al., 1987). The divergently transcribed *yp1* and *yp2* genes are located at 8F-9A and separated by a shared intergenic region of 1226 bp, whereas *yp3* is isolated and located at 12 BC (Figure 1.6). *yp1* and *yp2* each contain one intron whereas *yp3* has two. The *yp1* and *yp2* introns are 76 and 69 nucleotides in length respectively, and the two *yp3* introns are 62 and 72 nucleotides long. The first intron of *yp3* is located at a similar position to those of *yp1* and *yp2*, with their locations starting at +279 in *yp1*, +285 in *yp2* and +265 in *yp3*, relative to the transcriptional start site. The second intron of *yp3* is located at +710. Single transcripts are detected for *yp1* and *yp3* which are estimated to be 1.60 kb and 1.54 kb respectively by Northern blot experiments and RNA protection studies. Two transcripts are detected for *yp2* which differ at their 3' ends and are estimated to be 1.60 kb and 1.67 kb (Hung et al., 1982). The mature polypeptide sizes are 47 kD, 46 kD and 45 kD for *yp1*, *yp2* and *yp3* respectively, as determined by SDS-polyacrylamide gel electrophoresis. There is a high level of sequence similarity between the nucleotide sequences of these genes which is greatest in the 3' translated

Figure 1.6

Genomic arrangement of the three *yp* genes. The location on the X chromosome is indicated.

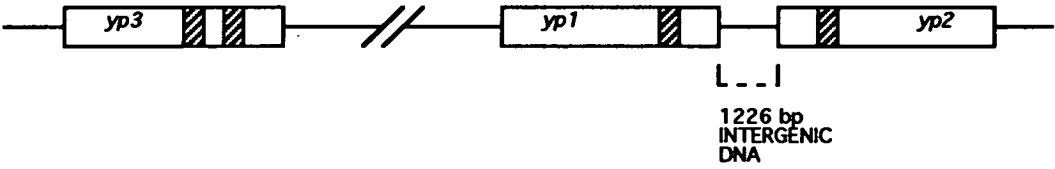


Transcripts are represented by arrows, with the arrowhead representing the direction of transcription, and spaces the location of the introns. The molecular weights of the mature polypeptides are given

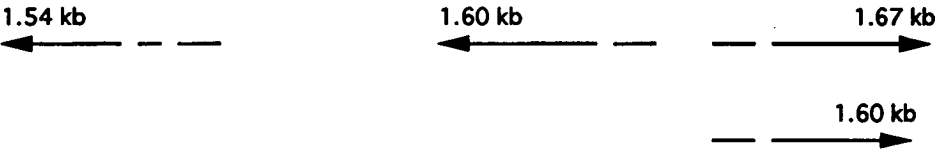
LOCATION ON  
X CHROMOSOME

12BC

8F - 9A



TRANSCRIPTS



PROTEIN  
PRODUCTS

45kD

47kD

46kD

regions. An identity of 53 % exists between the *yp1* and *yp2* genes and between *yp1* and *yp3*, and there is 48 % identity between *yp2* and *yp3*. This is reflected by the overall sequence identity of 43 % between the derived protein sequences, where the C-terminal sequences have 57 % identity compared to 29 % for the N-terminal regions (Garabedian et al., 1987). Due to the similarities in gene structure and the observed sequence identities, it is thought that these genes may have arisen from duplication events during evolution (Hatzopoulos and Kambysellis, 1987; Yan et al., 1987).

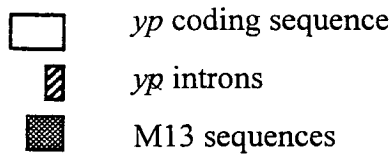
### 1.5.1 *yp1* and *yp2* Cis-acting Sequences

As well as their structural similarities, expression of the *yp* genes is coordinately regulated such that transcription occurs only in the fat body of adult females and in the ovarian follicle cells surrounding the developing oocyte during stages 8-10 of oogenesis. The mechanisms involved in this regulation must therefore involve the interaction of *cis*-acting elements with sex- and tissue-specific *trans*-acting factors present at the correct developmental stages for gene activation. Most of the investigations into *yp* gene control have focused on *yp1* and *yp2* and several *cis*-acting elements have been identified that are necessary for their correct pattern of expression. Various *yp1* and *yp2* gene constructs were tested *in vivo* using P-element mediated germ-line transformation of the *yp* genes and flanking sequences either tagged with M13 bacteriophage sequences, to distinguish them from native *yps* in Northern analysis, or fused to reporter genes such as alcohol dehydrogenase (*Adh*) or  $\beta$ -galactosidase (*lacZ*) (Figure 1.7).

Initial transformation experiments were carried out with a 5.0 kb genomic region containing a 3' truncated *yp1* gene, the intergenic region and the complete *yp2* coding sequence (Tamura et al., 1985). The truncated *yp1* transcripts could be distinguished from those of the endogenous gene in Northern analysis, and *yp2* transcription was assessed in a *yp2*<sup>-</sup> background. This demonstrated that the 5.0 kb region was sufficient for the correct sex-, temporal- and tissue-specific expression of both *yp* genes (Figure 1.7B). Garabedian et al. (1985) separated *yp1*

Figure 1.7

The *yp1* and *yp2* constructs tested for fat body enhancer elements



Arrows represent transcripts and the arrowhead the direction of transcription. Nucleotide positions are given in relation to the *yp1* transcription start site (+1). The ability of different constructs to direct *yp* transcription in the fat body were tested, using a number of techniques to distinguish the construct DNAs from the endogenous *yps*. A truncated *yp1* gene (B), levels of *yp2* transcripts in a *yp2*<sup>-</sup> background (B and O) and tagging the genes with M13 DNA inserts (C-F) allowed transcripts from these constructs to be distinguished from endogenous *yp* transcripts in Northern analysis. Reporter gene expression was assessed by staining for  $\beta$ -galactosidase activity in flies transformed with constructs containing a *hsp70-lacZ* gene fusion (G-M). *Adh* activity was assessed in flies transformed with construct N. The ability of each construct to direct ovarian or fat body transcription in adult females is indicated

- + - gene expression
- (+) - weak gene expression
- - no gene expression

		OVARIAN EXPRESSION	FEMALE FAT BODY EXPRESSION
A		+	+
B		+	+
C		-	+
D		+	-
E		-	+
F		-	-
G		-	+
H		-	+
I		-	-
J		-	+
K		-	+
L		-	(+)
M		-	(+)
N		-	+
O		+	+

and *yp2* at a *HindIII* site in the intergenic region such that *yp1* retained 886 bp of 5' flanking DNA and *yp2* retained 343 bp of upstream sequence. These constructs were tagged by inserting bacteriophage M13 DNA into the second exons (Figure 1.7C and D) such that their expression patterns in transformed flies could be distinguished from those of the endogenous *yp* genes by Northern analysis, using probes that contained the M13 sequences. Transcription of both construct occurred in the correct sex and developmental stage of the fly. However, *yp1* transcripts were detected only in the fat bodies of adult females whereas *yp2* transcripts were detected only in the ovaries, demonstrating that at least two different *cis*-acting elements had been separated that were necessary for the normal pattern of tissue-specific expression in these two different tissues.

### **1.5.2 Identification of Fat Body *Cis*-acting Elements Regulating *yp1* and *yp2* Gene Expression**

The region upstream of the M13 tagged *yp1* gene was investigated further to localise more precisely the *cis*-acting sequences required for fat body-specific expression of this gene by initially constructing a series of 5' deletions (Garabedian et al., 1986). A construct containing 321 bp of 5' flanking sequence (Figure 1.7E) was sufficient to direct the sex-, stage- and fat body-specific expression of *yp1*, but no transcripts were detected in the fat body of females transformed with a construct containing 89 bp of upstream DNA (Figure 1.7F). The fat body enhancer region was therefore localised to the region between nucleotides -321 and -89 upstream of the *yp1* transcription start site. A second set of deletions were constructed to determine more precisely the 3' boundary of this sequence (Figure 1.7G, H and I). These constructs were placed upstream of a *hsp70-lacZ* reporter gene such that gene activity could be detected by histochemical staining for  $\beta$ -galactosidase activity. A 3' deletion to -196 bp upstream of *yp1* contained sufficient DNA to direct reporter gene expression in the fat body of adult females, and these experiments also demonstrated that gene expression could be directed from a heterologous promoter. From these analyses it was concluded that 125 bp of

sequence located between nucleotide positions -196 and -321 upstream of the *yp1* transcription start site were sufficient for the correct pattern of *yp* expression in the fat body. This region was therefore termed the fat body enhancer (FBE). The FBE was tested independently by placing it in both orientations upstream of the *hsp70-lacZ* reporter gene (Figure 1.7J and K). The 125 bp sequence, with some neighbouring DNA (-886 to -196), was also tested in both orientations at a position 8.0 kb downstream of the *hsp70-lacZ* fusion construct (Figure 1.7L and M). In all instances, the 125 bp FBE was able to direct fat body-specific gene expression in adult females although the strength of the influence varied with the position and orientation of the sequence. The FBE therefore had a positive effect on gene transcription and displayed several of the characteristics used to define enhancer sequences, namely that they can exert their effects in either orientation and at varying distances from their target promoter (Serfling et al., 1985).

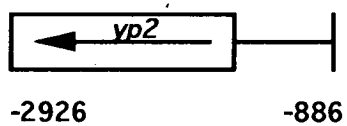
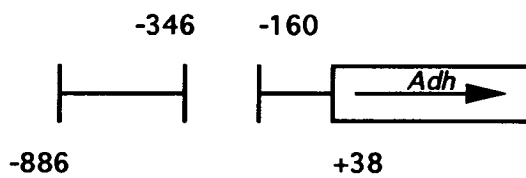
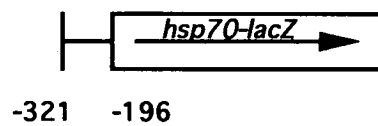
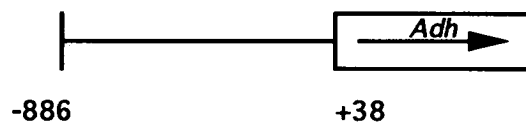
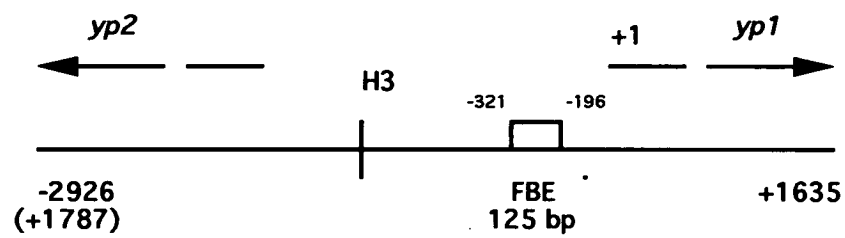
Although Garabedian et al. (1986) demonstrated that the 125 bp FBE was sufficient to direct the sex-, stage- and fat body-specific expression of gene transcription, it has since been demonstrated that this is not the only sequence in the intergenic region between *yp1* and *yp2* able to do so (Abrahamsen et al., 1993). Sequences upstream of *yp1* were fused to an *Adh* reporter gene and internal deletions were made instead of either 5' or 3' deletions. Using this approach, it was demonstrated that the FBE could be deleted entirely without losing fat body-specific expression in adult females (Figure 1.7N). The authors also demonstrated that a construct containing the entire coding region of *yp2* with 343 bp of upstream sequence was expressed in the fat body of adult females and secreted into the haemolymph of *yp2*<sup>-</sup> flies (Figure 1.7O). This construct also lacked the FBE and had previously been reported to confer ovarian, but not fat body, expression on the *yp2* gene (Garabedian et al., 1985).

The FBE is therefore sufficient but not essential for *yp* fat body expression, since other regions of the intergenic spacer are also capable of directing the correct sex-, stage- and fat body-specific pattern of expression either of the *yps* or reporter genes. These regions are summarised in Figure 1.8. The regulation of fat body-specific gene expression of *yp1* and *yp2* by *cis*-acting elements is therefore not



### Figure 1.8

Summary of the constructs able to direct the correct fat body-specific pattern of *yp* expression. Arrows represent transcripts and the direction of transcription. The location of the 125 bp fat body enhancer (FBE) identified by Garabedian et al. (1986) is illustrated. Nucleotide positions are given in relation to the *ypI* transcription start site (+1)



as straightforward as it initially appeared since there appears to be a redundancy of information in the intergenic region between *yp1* and *yp2*, with several regions able to confer the correct fat body-specific expression of these genes. However, the recent identification of *trans*-acting regulatory proteins that bind to the FBE provides further insight into the transcriptional regulation of *yp1* and *yp2* in this tissue.

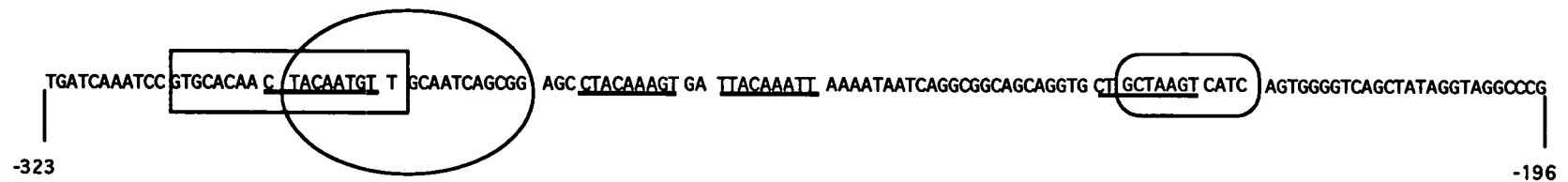
### 1.5.3 *Trans*-acting Factors that Bind to the *yp1/yp2* Fat Body Enhancer




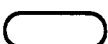
Although the FBE is not the only region in the *yp1/yp2* intergenic spacer containing sequences sufficient for fat body-specific gene expression in adult females, it has been successfully used in footprinting assays to identify several transcriptional regulatory proteins implicated in the control of fat body-specific *yp* expression. These are box B-binding factor-2 (BBF-2) which acts as a transcriptional activator, the CCAAT/enhancer-binding protein (C/EBP) and the adult enhancer factor-1 (AEF-1). These proteins were initially found to interact with the *alcohol dehydrogenase* (*Adh*) adult enhancer in *Drosophila melanogaster* to control expression of this gene in the adult fat body, and were subsequently shown by footprinting assays to have binding sites within the *yp* FBE (Abel et al., 1992; Falb and Maniatis, 1992a, 1992b). C/EBP and AEF-1 were shown to have overlapping recognition sites that are thought to be competitively bound, since AEF-1 can inhibit binding by C/EBP and thereby repress its activating potential (Falb and Maniatis, 1992b). The arrangement of these protein binding sites in the FBE is illustrated in Figure 1.9.

The male- and female-specific doublesex proteins (DSX<sup>M</sup> and DSX<sup>F</sup>) are also known to be involved in the regulation of *yp* expression and they have been shown to bind to four sequences in the 125 bp FBE (Burtis et al., 1991). These binding sites are included in Figure 1.9, and it can be seen that two of them overlap with the binding sites of the identified tissue-specific factors. The proximity and overlapping binding sites for the fat body binding factors with the DSX binding sites may explain why the sex- and fat body-specificity of gene expression have never been separated in the various constructs used to identify the FBE. This observation

### Figure 1.9

The *yp1/yp2* FBE (Garabedian et al., 1986) illustrating the location of the four DSX binding sites (Burtis et al., 1991) and the target sequences for the transcriptional regulatory proteins AEF-1, C/EBP and BBF-2 (Abel et al., 1992; Falb and Maniatis, 1992a, 1992b) identified by *in vitro* footprinting assays. Nucleotide positions are given relative to the *yp1* transcription start site



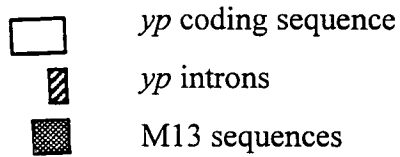
- 
 DSX binding sites
- 
 AEF-1 binding site (-295 to -312)
- 
 C/EBP binding site (-284 to -303)
- 
 BBF-2 binding site (-222 to -232)

also suggests that the sex- and tissue-specific proteins could be involved in the formation of a transcription initiation complex required to direct fat body-specific expression of the *yps* in adult females. Evidence to support this comes from recent mutational studies with the DSX binding sites in the FBE, the outcome of which suggested that DSX<sup>F</sup> had a positive influence on *yp* transcription in females (Coshigano and Wensink, 1993). In males, DSX<sup>M</sup> may prevent the binding of other *trans*-acting factors required to activate *yp* gene expression. Although DSX<sup>M</sup> and DSX<sup>F</sup> have been shown to bind to the same sequences in the *yp1/yp2* FBE *in vitro* (Burtis et al., 1991), the effects of DSX<sup>F</sup> *in vivo* could be altered by another sex-specific protein, an ideal candidate for which would be the protein product of the *ix* gene in the sex-determination gene hierarchy. However, the mechanisms involved in regulating the sex- and tissue-specificity of *yp* transcription remain to be elucidated. *In vitro* binding studies are currently underway to investigate the competitive interactions that occur between AEF-1, C/EBP, BBF-2 and the DSX proteins for binding to the FBE, and should provide further insight into the mechanisms involved in regulating the sex- and tissue-specific transcription of the *yps*.

Computer searches for the DSX binding sites and the tissue-specific *trans*-acting factors have revealed that other putative binding sites exist around the *yp* genes (Abrahamsen et al., 1993; Bownes et al., 1993) although they do not occur in such close proximity or with the overlapping organisation found in the FBE. However, *in vivo* analysis using reporter constructs with the FBE deleted from the *yp1/yp2* intergenic spacer have shown that the correct sex- and tissue-specific pattern of *yp* expression is still observed, suggesting that binding sites for the regulatory proteins in close proximity is not essential to confer female fat body-specific expression of the *yp* genes.

Figure 1.10

The *yp1* and *yp2* constructs tested for ovarian enhancer elements

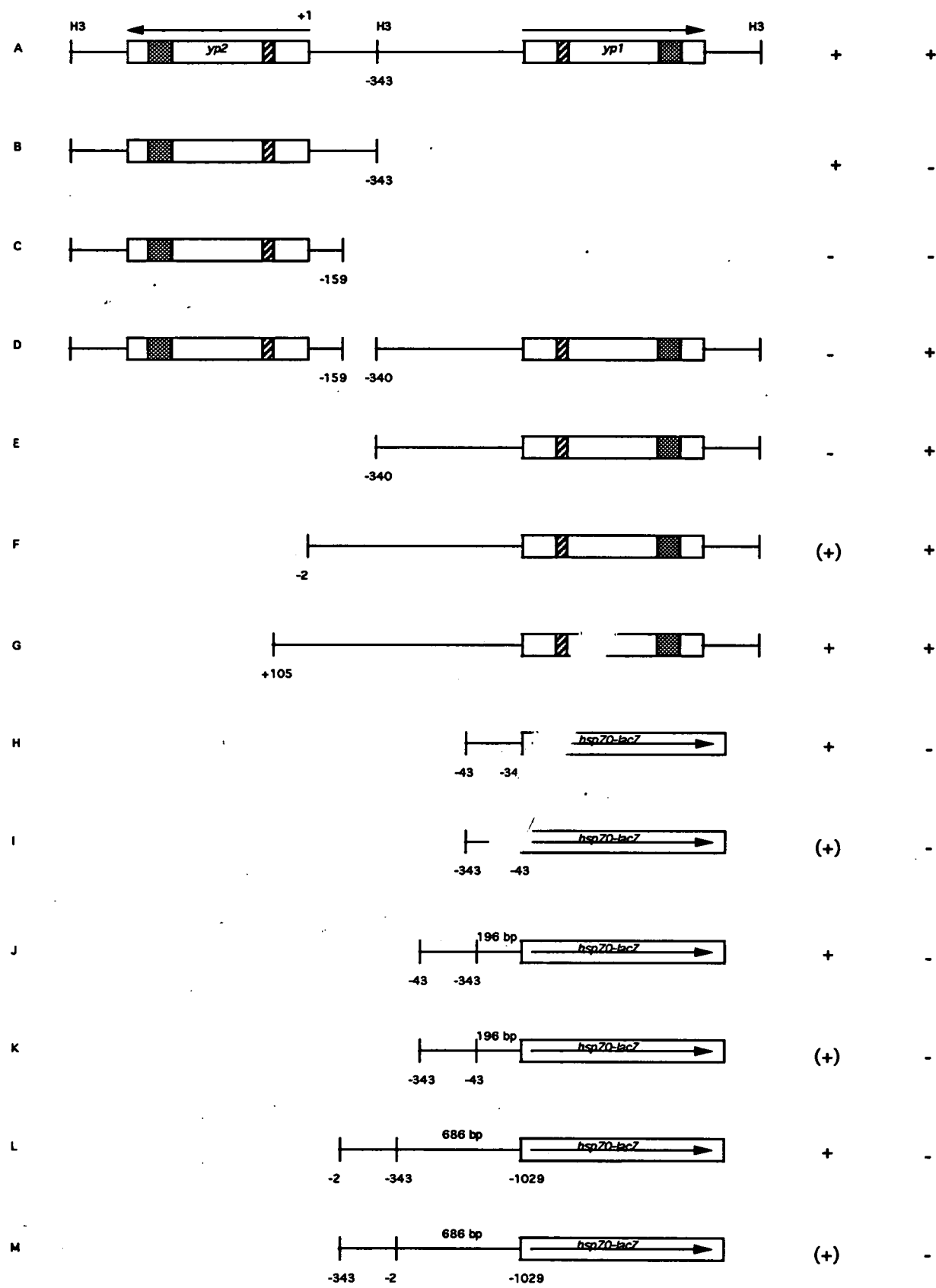


Arrows represent transcripts and the direction of transcription. Nucleotide positions are given in relation to the *yp2* transcription start site (+1). The ability of different constructs to direct *yp* transcription in the ovaries was tested. M13 DNA inserts were used to tag the *yp* genes, allowing transcripts to be distinguished from those of the endogenous *yps* in Northern analysis (A-G). Reporter gene expression was assessed by staining for  $\beta$ -galactosidase activity in flies transformed with constructs containing a *hsp70-lacZ* gene fusion (H-M). The ability of each construct to direct ovarian or fat body transcription in adult females is indicated

- + - gene expression
- (+) - weak gene expression
- - no gene expression

OVARIAN  
EXPRESSION

FEMALE  
FAT BODY  
EXPRESSION





#### 1.5.4 Identification of Ovarian *Cis*-acting Elements Regulating *yp1* and *yp2* Gene Expression

Initial transformation experiments that separated *yp1* and *yp2* at the *HindIII* site in the intergenic spacer showed that M13 tagged *yp2* transcripts were only present in ovarian tissue (Garabedian et al., 1985). The *yp2* gene was contained within a 2.8 kb DNA fragment that included 343 bp of upstream sequence. Deletions made in this upstream region revealed that a 184 bp sequence located between nucleotides -159 to -343 upstream of the *yp2* transcription start site was sufficient for the ovarian transcription of this gene (Figure 1.10B and C) (Logan et al., 1989). This fragment was shown to have a positive tissue-specific influence on both *yp1* and *yp2* expression, as deleting it from this gene pair in their normal arrangement caused a reduction in their transcript levels in the ovary (Figure 1.10D). However, it was also observed that the reduction in ovarian *yp2* transcript levels with this *yp1/yp2* construct was not as great as that which occurred when the *yp2* gene was tested independently. This suggested the presence of a second ovarian enhancer element which was located by assessing the ovarian levels of M13 tagged *yp1* transcripts with constructs containing varying amounts of upstream sequence. A *yp1* construct truncated at -343 bp upstream of the *yp2* cap site (and therefore containing 886 bp of upstream sequence) had already been shown not to express *yp1* RNA in the ovaries (Figure 1.10E). Surprisingly, when the *yp1* 5' flanking region was extended to the *yp2* cap site such that it contained the entire intergenic region, only very low levels of ovarian transcripts were detected, although this included the 184 bp region that gave high levels of ovarian *yp2* transcription (Figure 1.10F). Although this 184 bp fragment was also shown to have a positive tissue-specific influence on the ovarian transcription of both *yp1* and *yp2* in their normal genomic arrangement, it was obviously not sufficient on its own to confer normal levels of *yp1* ovarian expression. A second ovarian enhancer element (OE2) was found to be contained within a 105 bp sequence located in the first exon of the *yp2* coding sequence (Figure 1.10G). When this was added to the *yp1* gene and intergenic region, transcript levels were restored to those observed with the *yp1/yp2* gene pair.

OE2 therefore appeared to act across the *yp2* promoter and the intergenic region to enhance levels of ovarian *yp1* transcripts.

The *yp2* upstream sequence spanning -43 to -343, containing the 184 bp ovarian enhancer sequence first identified, was investigated further by placing it at various distances in both orientations upstream of a *hsp70-lacZ* reporter gene (Figure 1.10H-M). This allowed the developmental profile within the ovary, and the cell type specificity of reporter gene expression to be observed in more detail. When this 301 bp construct was placed in the correct orientation (with respect to *yp2*) at distances of 196 bp and 686 bp upstream of the reporter gene (Figure 1.10J and L) the developmental- and cell type-specific pattern of *yp* ovarian expression was not affected and this region was subsequently termed ovarian enhancer 1 (OE1). However, when the orientation of OE1 was reversed at these locations (Figure 1.10K and M) both the stage-specificity and cell type-specificity of expression varied between different transformed lines. The reason for this was unclear but may have been due to the presence of *yp2* promoter remnants interfering with enhancer activity or an increased sensitivity to nearby regulatory elements at the different chromosomal sites of insertion. Nevertheless, when in its normal orientation, OE1 consistently directed the ovarian *yolk protein* expression pattern on a *lacZ* reporter gene from a heterologous promoter (Logan et al., 1989).

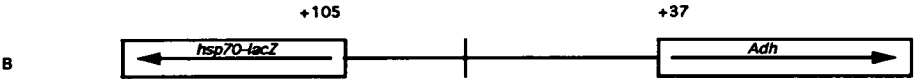
It was later shown that different sub-fragments of OE1 were responsible for directing the cell type-specificity of *yp* expression in the follicle cells of developing egg chambers (Logan and Wensink, 1990). Several types of follicle cells are present during specific stages of oogenesis, namely the columnar cells, the columnar main body cells, the anterior/posterior pole cells, the border cells and cells covering the nurse cells (Figure 1.3A and B). By linking the native *yp1* and *yp2* genes to *Adh* and *lacZ* reporter genes respectively it was initially demonstrated that these genes were transcribed in only the first three cell types mentioned above (Figure 1.11B). A similar cell type-specific pattern of expression was observed when OE1 itself was placed upstream of a *hsp70-lacZ* reporter gene. The only difference noted was that OE1 caused faint staining of the border cells, which did not occur with the *yp1/yp2* construct (Figure 1.11C). Removal of 91 bp from the 3' end

### Figure 1.11

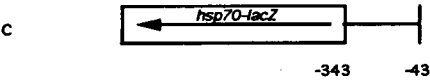
Constructs used to assess the effects of OE1 sub-fragments on reporter gene expression in the follicle cells of developing egg chambers (adapted from Logan and Wensink, 1990). The genomic arrangement of *yp1* and *yp2* is shown in the top figure (A), which were replaced by the *Adh* and *lacZ* reporter genes respectively (B). All nucleotide positions are given relative to the *yp2* transcription start site (except +37, which is relative to the *yp1* cap site). The orientation of reporter gene transcription is shown by the arrows. The expression patterns in columnar main body cells (CMB), border cells (BC) and anterior pole cells (AP), assessed by  $\beta$ -galactosidase staining intensity, are indicated

+++ - very strong  
+ - very faint  
- - undetectable

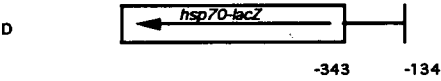
CMB      BC      AP



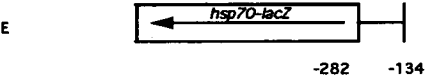
+++      -      +++



+++      +      +++



-      +++      +++



-      +++      +++

of OE1 (Figure 1.11D) resulted in a loss of reporter gene expression in the columnar main body cells, which are a major class of cells that normally express the *yp* genes. Staining of the border cells was also much stronger than observed with lines transformed with the entire OE1 region. Removing a further 61 bp from the 5' end of the truncated OE1 did not alter these observations (Figure 1.11E). Expression remained strong in the anterior pole cells and border cells but was not seen in the columnar main body cells. From this it was concluded that the central 149 bp of OE1 could direct transcription in the anterior pole cells and border cells, but 91 bp of 3' sequence was required to direct expression in the columnar main body cells and to suppress expression in the border cells. A combination of the effects of both positively and negatively acting *cis*-regulatory elements therefore appears to determine the cell type-specificity of *yp* expression in the ovarian follicle cells. A similar regulatory system appears to be involved in controlling the temporal specificity of expression of the *Drosophila chorion* and *vitelline membrane* genes.

The chorion genes comprise another family of genes that are expressed in the ovarian follicle cells, once YP synthesis has ceased, during stages 10b to 14 of oogenesis. The chorion genes are under strict temporal control and each gene has a precise and limited period of expression. For example, transcripts from the *s36* developmentally "early" gene are detected at stages 10b to 12, and disappear at stage 13 (Parks et al., 1986) whereas the *s15* "late" gene transcripts start to appear during stage 13 of oogenesis and are most abundant at stage 14 (Griffin-Shea et al., 1982). Analysis of *s15* 5' flanking DNA revealed that the precise temporal specificity of gene expression resulted from the combined effects of closely linked positively and negatively acting *cis*-regulatory elements (Mariani et al., 1988). Similarly, an analysis of the upstream region flanking the *s36* gene revealed that multiple, spatially specific *cis*-regulatory elements determined the spatial pattern of chorion gene expression in the anterior and posterior poles of the follicular epithelium (Tolias and Kafatos, 1990). More recently, Bienz-Tadmor et al. (1992) reported that the tissue-specificity of chorion gene expression is mediated by *cis*-regulatory regions that restrict expression of the chorion genes in other tissues. The consistent element in these reports is that the observed pattern of chorion gene

expression is due to the combined effects of *cis*-regulatory positive acting elements that activate gene expression, and negative acting elements that prevent this gene activation in inappropriate tissues or at the wrong developmental stage.

The *vitelline membrane* genes are only expressed in the columnar follicle cells of stage 8, 9 and 10 egg chambers (Gigliotti et al., 1989; Savant and Waring, 1989). Four different vitelline membrane genes have been identified, namely *VM26A1* (also called *Tu-2*), *VM26A2* (also called *Sv23* and *Tu-4*), *VM34C* and *VM32E* (also called *VMP32E*) (Higgins et al., 1984; Mindrinos et al., 1985; Burke et al., 1987; Gigliotti et al., 1989). Analysis of the 5' flanking region of *VM26A1*, with fragments fused upstream of *hsp70-lacZ* reporter genes, revealed that a 176 bp region between -352 and -176 bp upstream of the transcription start site was capable of promoting the vitelline membrane expression pattern. However, sequences flanking this (-617 to -299 and -158 to -33) were required for full levels of reporter gene expression, although they could not direct *lacZ* reporter gene expression when tested independently. In addition to this, abnormal border cell expression was observed when the 176 bp region was tested independently, which was suppressed by an adjacent negative control element in the sequence spanning nucleotides -175 to -99 (Jin and Petri, 1993).

It would not be surprising if the mechanisms involved in the regulation of *yolk protein* gene expression resembled those of the *vitelline membrane* and *chorion* genes, as these families of genes have overlapping expression in the ovarian follicle cells and are all transcribed with a highly specific developmental profile.

### **1.5.5 Ovarian *Trans*-acting Factors**

There have been no reports on the isolation of regulatory proteins that could be involved in governing the tissue-specific ovarian expression of the divergently transcribed *yp1* and *yp2* genes. This has possibly not been pursued due to the complexities that arise from having shared promoter sequences in the

intergenic region and the known requirement for two *cis*-acting enhancers, OE1 and OE2, for correct ovarian *yp* expression.

It is interesting to note that the transcription factors AEF-1, C/EBP and BBF-2 that have been shown to bind to the FBE (and are therefore thought to be involved in regulating the fat body-specific expression of *yp1* and *yp2*) are present in ovarian nuclear extracts. These factors were initially found to interact with the *Adh* adult enhancer to control expression of this gene in the adult fat body of *Drosophila melanogaster* and, like the *yps*, the ovaries are the other adult tissue where *Adh* is expressed (Abel et al., 1992; Falb and Maniatis, 1992a, 1992b). Both *Adh* and the *yp* genes are therefore specifically expressed in the adult fat body and ovaries, and the factors regulating their fat body-specific expression are also present in ovarian nuclear extracts. Although there is little experimental evidence to indicate that the factors involved in regulating fat body-specific transcription of these genes could also regulate their expression in the ovaries, a combination of factors will probably regulate expression in each tissue and it is possible that some of the factors may be present in the fat body and ovaries. However, this is entirely speculative as it has not been demonstrated that these factors are necessary for adult *Adh* expression in the ovaries, and it is not known if *Adh* is expressed in the ovarian follicle cells. Also, there is no evidence demonstrating that these factors actually have a role in the regulation of *yp* transcription *in vivo* as binding to the FBE was demonstrated using *in vitro* DNaseI footprinting assays. Other factors known to influence *yp* transcription in the fat body have not been required to regulate expression in the ovary. For example the hormonal regulation of *yolk protein* transcription in these two tissues is different (see sections 1.4.2.1 and 1.4.2.2.), and only fat body expression of the *yps* is dependent upon the continuous action of the somatic sex-determination pathway i.e. regulated by DSX (see section 1.4.3.1). Whether AEF-1, C/EBP and BBF-2 are involved in regulating the ovarian expression of the *yp* genes would require an assessment of *yp* expression in the ovaries of *Drosophila* that are mutant for the genes encoding these transcription factors. Preliminary studies of this nature have been carried out for C/EBP, as the *Drosophila* homologue of this protein has been identified and is encoded by the *slow border (slbo)* gene (Montell et al.,

1992). YP synthesis in homozygous mutants appears to be at a similar level to that observed in wild type flies (M.Bownes, personal communication). Preliminary evidence therefore suggests that C/EBP will not be involved in regulating tissue-specific expression of the *yps* in the ovary. However, once this type of analysis is carried out in more detail, and can be performed with the AEF-1 and BBF-2 transcription factors (for which there are no mutants at present) it should become clear whether these proteins have a role in regulating ovarian *yp* expression. Until then, binding studies with these transcription factors and the identified ovarian enhancers, OE1 and OE2, could be performed to provide some indication as to whether such interactions are possible.

Another protein, yolk protein factor 1 (YPF1), isolated from *Drosophila* KcO tissue culture cells (an embryonic cell line derived from first instar larvae) and also present in ovarian and early embryonic extracts, was identified several years ago. This bound with high affinity to a 31 bp sequence located 148 bp downstream of the *yp1* transcription start site and was therefore thought to be potentially involved in regulating an aspect of the tissue-specific pattern of *yp* expression (Mitsis and Wensink, 1989a, 1989b). Computer sequence comparisons located a 26 bp sequence with 77 % sequence identity to the 31 bp sequence at a similar position 140 bp downstream of the *yp3* transcription start site. However, since this was in the coding region of a gene family the similarity may not necessarily have been for YPF1 binding, but rather required for the structure or function of the yolk proteins themselves. Deletion of this binding site from *yp1* did reduce the levels of *yp1* transcripts in females (Mitsis and Wensink, 1989a) although this could have been due to transcript instability rather than a loss of binding by a transcriptional activator protein. More recently, YPF1 has been reported to be a heterodimeric DNA-binding protein and the *Drosophila* homologue of Ku, a human DNA-dependent protein kinase (Jacoby and Wensink, 1994). This is thought to be involved in the initiation of gene transcription by phosphorylating the carboxyl-terminal domain of RNA polymerase II. This induces a conformational change such that pol II switches from being a promoter binding protein to become part of a transcription elongation complex. It has been proposed that the association of YPF1



with *yp* sequences close to the transcription start site could position the equivalent kinase subunit adjacent to a transcription initiation complex which would facilitate high levels of *yp* transcription during oogenesis. The profile of YPF1 transcripts is consistent with this proposed function as it is ubiquitously expressed at low levels throughout *Drosophila* development, with the highest transcript levels being in the oocyte and early embryo. However, this does not implicate YPF1 in the control of tissue-specific *yp* transcription in the ovary, and the factors regulating this are yet to be isolated.

### 1.5.6 *yp3* Cis-acting Sequences

Although the expression of *yp1* and *yp2* have been investigated in some detail, a detailed analysis of *yp3* has not been carried out although this gene has also been cloned and sequenced (Barnett et al., 1980; Garabedian et al., 1987). The *yp3* gene was isolated on a 4.28 kb fragment that contained 1821 bp of upstream sequence and 825 bp of 3' flanking DNA (Garabedian et al., 1987). A preliminary study to locate regulatory sequences governing *yp3* expression determined that the *yp3* coding region with 704 bp of sequence upstream from the transcription start site and 825 bp of downstream DNA was expressed in the correct tissues of females when transformed into a YP3<sup>-</sup> background (Liddell and Bownes, 1991). The constructs used in this analysis are given in Figure 3.2A and B (Chapter 3).

Since all three *yolk protein* genes are coordinately transcribed with a high degree of specificity, the factors involved in the regulation of *yp1* and *yp2* expression are expected to be similar, if not the same, for *yp3*. Other putative regions of importance in the control of *yp3* expression have therefore been derived from sequence comparisons between the enhancer regions identified for the tissue-specific expression of *yp1* and *yp2*. Garabedian et al. (1987) found two regions of sequence similarity between the 125 bp FBE and *yp3* flanking sequences. One was located at nucleotide positions -909 to -924 upstream of the *yp3* transcription start site and had 87 % identity to the sequence located between nucleotides -254 to -269 of the FBE. The other region was located downstream of the *yp3* coding sequence

between nucleotide positions +2081 to +2107 and had 70 % identity to the sequence located between nucleotides -283 to -309 of the FBE (Figure 1.12A and B).

Computer sequence comparisons were also carried out between the two ovarian enhancers, OE1 and OE2, since they were both involved in directing similar cell-type specificities of ovarian *yolk protein* gene expression (Logan and Wensink, 1990). A region of sequence similarity was found to exist not only between OE1 and OE2, but between two *yp3* sequences and an ecdysone response element from the *Drosophila hsp27* gene (Figure 1.12D). A 9 bp sequence in OE1 (-119 to -127) had a match of 7/9 nucleotides to a sequence in OE2 (+50 to +58) and 9/9 nucleotides to a sequence in *yp3* (-632 to -640). Another 12 bp sequence adjacent to this in OE1 (-105 to -116) had a match of 11 bp to a different sequence in *yp3* (-1566 to -1577). Since *yp3* has the same pattern of expression as *yp1* and *yp2* it is possible that these sequences could be involved in directing the ovarian expression of this gene, although this would have to be confirmed by making deletions in these regions followed by *in vivo* analysis of reporter gene expression. Interestingly, the regions of sequence identity between OE1, OE2 and *yp3* also had sequence similarity to an ecdysone response element which could be related to the influence of 20-hydroxyecdysone on *yolk protein* gene expression. Again, this would have to be examined further by mutational analysis and the *in vivo* effects, to determine the significance of this finding.

A region of sequence similarity was also reported to exist between nucleotides +71 to +102 of OE2, located in the coding sequence of *yp2*, and a similar position, between nucleotides +74 to +108, in the coding sequence of *yp3* (Figure 1.12C, Liddell and Bownes, 1991). Although sequence comparisons can provide an indication of DNA sequences that may be of functional importance, the significance of such sequence identities, and whether they are actually involved in regulating *yp3* expression, should be determined experimentally.

## Figure 1.12

**A and B**      Regions of sequence similarity, identified by computer sequence comparisons, between the *yp1/yp2* FBE and *yp3* flanking sequences (Garabedian et al., 1987). Nucleotide positions associated with the FBE are relative to the *yp1* transcription start site, and *yp3* positions relative to the transcription start site of *yp3*. Vertical lines indicate nucleotide matches. Percentage sequence identities are 87 % in **A**, and 70 % in **B**

**C**              Sequence similarity identified by computer sequence comparisons to exist between OE2, located in the coding region of *yp2*, and a sequence located in a similar position downstream of the *yp3* cap site (Liddell and Bownes, 1991), with 71 % sequence identity

**D**              Regions of sequence similarity between OE1, OE2, *yp3* and the ecdysone response element of the *Drosophila hsp27* gene

A.      FBE      -254      TTATTTTAATTTGTAA      -269  
              *yp3*      -909      TTATTTTACTATGTAA      -924

B.      FBE      -283      GCACAAC TACAATGTTGCAATCAGCGG      -309  
              *yp3*      +2081      GCACAACGAGATGGATGCATTGAGTGG      +2107

C.      OE2      +102      GGCGACCGCCAGAAGGCA ... GGCCATAACGCAAA      +71  
              *yp3*      +108      GGCCGCCACCAGGAGGCAGGTGGCCAGCAGGCAAA      +74

D.      *yp3*      TTTGTTCTGATC      (-1566 to -1577)  
              *yp3*      GAATCAATG      (-632 to -640)  
              OE1      TGACGCTGCGAATCAATGCATTTGTGCTGATC      (-134 to -105)  
              EcRE      GAGACAAGGGTTCAATGCACTTGTCCAA  
              OE2      GCGCAGAGGATTCATTGTGGCT . TCCAA      (+66 to +40)

### 1.5.7 *yp3* Trans-acting Factors

Computer searches using consensus sequences for the binding sites of the factors implicated in the fat body-specific expression of *yp1* and *yp2*, namely the DSX proteins and the AEF-1, C/EBP and BBF-2 transcription factors, have identified putative binding sites in the *yp3* upstream region (Bownes et al., 1993). Interestingly, these do not occur with the same proximity or overlapping conformation seen in the 125 bp FBE, which again suggests that although this is sufficient to confer fat body-specific expression of the *yp1* and *yp2* genes, it may not be an essential prerequisite. Again, the significance of these findings can only be established following experimental *in vivo* analysis with various *yp3* gene constructs.

The aim of this PhD project was to carry out such an analysis of the *yp3* gene and its flanking sequences in order to identify the *cis*-acting DNA sequences required for the observed stage-, sex- and tissue-specific pattern of ovarian and fat body expression. Since *yp3* was located some distance away from *yp1* and *yp2* on the X chromosome and had been isolated with its own regulatory sequences, the *cis*-acting factors involved in the control of its transcription were anticipated to be more straightforward to determine than those for the divergently transcribed *yp1* and *yp2* genes with their shared regulatory elements. Once the enhancer regions had been located, the aim of the project would then be to try and identify the regulatory proteins interacting with them. Due to their highly coordinated pattern of expression, the factors involved in the transcriptional regulation of all three *yp* genes are expected to be the same, or highly similar. The results of this research could therefore be compared to those reported for *yp1* and *yp2* in an attempt to clarify the current knowledge on the transcriptional control of this small gene family.

## **CHAPTER 2**

### **Materials and Methods**

## **2.1 MEDIA**

### **2.1.1 Bacterial Media**

All media were made up in distilled water and sterilised by autoclaving (15psi/20mins). Ampicillin was added to media and plates to a final concentration of 100µg/ml when required.

Luria Broth	10g Difco Bacto Tryptone, 5g Difco Bacto Yeast Extract, 5g NaCl. Made up to 1 litre with H <sub>2</sub> O and adjusted to pH 7.2
Luria Agar	As Luria Broth with the addition of 15g Difco Bacto Agar per litre
2x TY Broth	16g Difco Bacto Tryptone, 10g Difco Bacto Yeast Extract, 5g NaCl. Made up to 1 litre with H <sub>2</sub> O and adjusted to pH 7.4
2x TY Agar	As 2x TY Broth with the addition of 15g Difco Bacto Agar per litre
5x Spizizen Salts	10g (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 20g K <sub>2</sub> HPO <sub>4</sub> , 30g KH <sub>2</sub> PO <sub>4</sub> , 5g tri-sodium citrate, 1g MgSO <sub>4</sub> . Made up to 1 litre with H <sub>2</sub> O
Minimal Medium	80ml 5x Spizizen salts, 4ml 20 % (w/v) glucose, 0.1ml Vitamin B1 (5mg/ml). Made up to 400ml with H <sub>2</sub> O
Minimal Agar	As minimal medium with the addition of 6g Difco Bacto Agar per 400ml

### 2.1.2 *Drosophila* Media

Cornmeal Food	250g Cornflour, 500g Sugar, 175g Yeast pellets, 100g Difco Bacto Agar. Made up to 10 litres with distilled water
‘French’ Fly Food	7.5g Oxoid No.3 Agar, 55g polenta, 55g dried flake yeast, 10mls nipagen (150mg/ml in 95 % ethanol), 100mls distilled water
Egg Collection Media	1.5g Difco Bacto Agar per 100mls pure apple juice

## 2.2 BACTERIAL, PLASMID AND *DROSOPHILA* STOCKS

All bacterial, plasmid and fly stocks used are given in tables 2.1 to 2.4

Table 2.1 Bacterial stocks

Cell Type	Genotype	Use	Reference
HB101	<i>supE44, ara14, galK2, lacY1, proA<sub>2</sub> rpsL20, xyl-5, mtl-1, recA13, Δ(mcrC-mrr), HsdS-(r-m-)</i>	Used as a general plasmid host	Maniatis et al., 1982
NM522	<i>supE, thi-1, Δ(lac-proAB), Δ(hsdSM-mcrB)5(rk-,mk-), [F', proAB,lacI<sup>f</sup> ZΔM15</i>	Used as a host for pBluescript clones. Selected for on minimal media	Gough and Murray, 1983



**Table 2.2**      Plasmids for subcloning

<b>Plasmid</b>	<b>Genotype</b>	<b>Comments</b>	<b>Reference</b>
pGEM-1/2	Amp <sup>r</sup>	Vector used for subcloning <i>Drosophila</i> <i>yp</i> sequences	Promega Biotechnology (Melton, 1984)
p-194.70Z	Amp <sup>r</sup>	Plasmid containing <i>hsp70</i> promoter fused to <i>E.coli lacZ</i> . Used to isolate <i>hsp70</i> sequences for construction of pERI vector	Xiao and Lis, 1988
pUC19	Amp <sup>r</sup>	Vector used for subcloning <i>Drosophila</i> <i>hsp70</i> promoter sequences	Messing, 1983
pBluescript SK+	Amp <sup>r</sup>	Vector used for subcloning <i>Drosophila</i> <i>hsp70</i> and <i>yp3</i> sequences	Stratagene

**Table 2.3** Plasmids for P-element mediated germ line transformation

Plasmid	Genotype	Comments	Reference
pCaSpeR-AUG- βgal	Amp <sup>r</sup>	P-element transformation vector containing <i>white</i> as the selectable marker. Used to test the effects of <i>yp3</i> promoter sequences on <i>lacZ</i> reporter gene expression	Thummel et al., 1988
pERI	Amp <sup>r</sup>	pCaSpeR-AUG-βgal transformation vector modified to contain <i>Drosophila hsp70</i> promoter sequences to test the effects of <i>yp3</i> constructs lacking the <i>yp3</i> promoter on <i>lacZ</i> reporter gene expression	
phs70Δ2-3wc	Amp <sup>r</sup>	Helper plasmid for P-element transformation containing a P-element, lacking the germline specifically spliced intron, transcribed from the <i>hsp70</i> promoter	Gift from Carol McLean, obtained from Jean-Maurice Dura

Table 2.4      *Drosophila* stocks

Name	Relevant Features	Reference
Oregon R (OrR)	A <i>Drosophila melanogaster</i> wild type strain	Lindsley and Grell, 1968
W <sup>K</sup>	A <i>white</i> <sup>-</sup> strain of <i>Drosophila</i> used for P-element mediated germ line transformation	Lüning, 1981

## 2.3 SOLUTIONS

Chemicals were obtained from Aldrich, BDH, Boehringer Mannheim, Fisons, NBL and Sigma.

Radioisotopes, Hybond-N and Hybond-N+ nylon membranes were obtained from Amersham.

Enzymes (restriction endonucleases, T4 DNA ligase, T4 DNA polymerase, Klenow fragment of DNA polymerase, DNase I, Rnase A, Proteinase K) were obtained from BRL, NBL and Pharmacia.

Standard solutions were made using sterile distilled water and sterile, baked glassware. Solutions were sterilised by autoclaving (15 psi/15 min) or by passing through a 0.45µm nitrocellulose filter. Solutions not described in the text are detailed below.

TE	10mM Tris-HCl pH 7.4-8.0, 1mM EDTA
----	------------------------------------

0.5M EDTA	0.5M Diaminoethanetetra-acetic Acid, pH 8.0
-----------	---

10x TBE	0.89M Tris, 0.89M Boric Acid, 10mM EDTA
10x DNA gel loading	100mM EDTA pH 8.0, 0.25 % (w/v) Bromophenol Blue, buffer 50 % (v/v) glycerol
Ringer's Solution	3.2g NaCl, 3.0g KCl, 1.8g MgSO <sub>4</sub> ·7H <sub>2</sub> O, 0.69g CaCl <sub>2</sub> ·2H <sub>2</sub> O, 1.79g Tricine, 3.6g glucose, 17.1g sucrose. Made up to 1 litre with H <sub>2</sub> O, adjusted to pH 6.95 with NaOH, filter sterilised and stored at 4°C

### 2.3.1 Sequencing Gel Solutions

6 % Sequencing Gel Mix	115g Urea in 150ml H <sub>2</sub> O - heated to dissolve, 14.25g Acrylamide, 0.75g N-M-methylene-bisacrylamide, 25ml 10x TBE. H <sub>2</sub> O added to 250ml and stored in dark at 4°C
Sequencing Gel Solution (45cm x 18cm x 0.2mm)	40ml 6 % sequencing gel mix, 60μl 25 % (w/v) ammonium persulphate, 60μl Temed
Sequencing Gel Solution (60cm x 18cm x 0.2mm)	45ml 6 % sequencing gel mix, 70μl 25 % (w/v) ammonium persulphate, 70μl Temed

### 2.3.2 Southern Blot Solutions

20 X SSC (Blotting Buffer)	3M NaCl, 0.3M Na <sub>3</sub> -Citrate pH 7.0
OLB (Oligo Labelling Buffer)	<u>Solution O</u> 0.125M MgCl <sub>2</sub> , 1.25M Tris-HCl pH 8.0 <u>Solution A</u>

0.95ml Solution O, 18ml  $\beta$ -Mercaptoethanol, 25 $\mu$ l 20mM dATP, 25 $\mu$ l 20mM dTTP, 25 $\mu$ l 20mM dGTP

#### Solution B

2M Hepes titrated to pH 6.6 with 4M NaOH

#### Solution C

Hexadeoxyribonucleotides (Pharmacia) in TE buffer at 90 OD units/ml

OLB is made by mixing solutions A, B and C at a ratio of 2:5:3 which is subsequently filter sterilised and stored at -20°C

### **2.3.3 Protein Gel Solutions**

2x Polypeptide Sample Buffer	50mM Tris-HCl pH 6.8, 1 % (w/v) SDS, 20 % (v/v) glycerol, 0.1 % (v/v) $\beta$ -mercaptoethanol, 0.01 % (w/v) Bromophenol Blue
Acrylamide stock	29.2 % (w/v) Acrylamide, 0.8 % (w/v) N-M-methylene-bisacrylamide
10 % Separating Gel Solution	5ml 3M Tris-HCl pH 8.8, 0.4ml 10 % (w/v) SDS, 13.8ml Acrylamide stock, 20.8ml H <sub>2</sub> O, 133 $\mu$ l 15 % (w/v) ammonium persulphate, 25 $\mu$ l Temed
Stacking Gel Solution	1.25ml 0.5M Tris-HCl pH 6.8, 100 $\mu$ l 10 % (w/v) SDS, 1.25 ml Acrylamide stock, 7.3ml H <sub>2</sub> O, 75 $\mu$ l 15 % (w/v) ammonium persulphate, 7 $\mu$ l Temed
10x Gel Running Buffer	1.5M Glycine, 0.2M Tris, 1 % (w/v) SDS

### 2.3.4 $\beta$ -galactosidase Staining Solutions

X-gal stock	8 % (w/v) 5-bromo-4-chloro-3-indonyl- $\beta$ -galactosidase, made up in dimethyl formamide and stored in a dark container at 4°C
X-gal Staining Buffer	<u>Solution A</u> 0.5M Na <sub>2</sub> HPO <sub>4</sub> /NaH <sub>2</sub> PO <sub>4</sub> <u>Solution B</u> 10mM Solution A, 150mM NaCl, 1mM MgCl <sub>2</sub> , 7mM Potassium Ferrocyanide, 7mM Potassium Ferricyanide. Stored in a dark container at 4°C. Immediately prior to use, X-gal is added to a concentration of 0.2 % (w/v)

### 2.3.5 Solutions for *lacZ* Assays

20-hydroxyecdysone	Obtained from Sigma and made up at a concentration of 10 <sup>-2</sup> M. 2.4mg dissolved in 500 $\mu$ l Ringer's Solution containing 2 % ethanol. Stored at -20°C for up to 6 months
Homogenisation Buffer	60mM Na <sub>2</sub> HPO <sub>4</sub> .7H <sub>2</sub> O, 40mM NaH <sub>2</sub> PO <sub>4</sub> .H <sub>2</sub> O, 10mM KCl, 1mM MgSO <sub>4</sub> .7H <sub>2</sub> O, 50mM $\beta$ -mercaptoethanol. Stored at 4°C for up to 1 month
ONPG	o-Nitrophenyl- $\beta$ -galactopyranoside (Sigma), at 4mg/ml in 100mM Na <sub>2</sub> HPO <sub>4</sub> .7H <sub>2</sub> O pH 7.0

### 2.3.6 Solutions for the Preparation of Ovarian Nuclear Extracts

MTBS	135mM NaCl, 5mM KCl, 1mM CaCl <sub>2</sub> pH 7.4, 0.5mM MgCl <sub>2</sub> , 1.5mM Na <sub>2</sub> HPO <sub>4</sub> ·7H <sub>2</sub> O, 2.5mM Tris-HCl pH7.5
PMSF	phenylmethanesulfonylfluoride (Sigma) freshly made up at 100mM in isopropyl alcohol
Solution I	10mM HEPES pH 7.9, 10mM KCl, 1.5mM MgCl <sub>2</sub> , 0.5mM dithiothreitol, 0.5mM PMSF
Solution II	10mM HEPES pH 7.9, 400mM NaCl, 1.5mM MgCl <sub>2</sub> , 0.5mM dithiothreitol, 0.5mM PMSF; 5% (v/v) glycerol

### 2.3.7 Gel Retardation Assay Solutions

Incubation Buffer	20mM Tris-HCl pH 8.0, 400 mM KCl, 1mM EGTA (filter sterile), 1mM EDTA pH 7.5, 2mM MgCl <sub>2</sub> pH 6.0, 1mM dithiothreitol, 500µg/ml BSA, 10 % (v/v) glycerol. Stored in 200µl aliquots at -20°C. Prior to use PMSF added to 100µg/ml, and Leupeptin and Pepstatin A added to 1µg/ml
5x Tris-glycine	30.28g Tris-base, 142.7g Glycine, 3.92g EDTA. H <sub>2</sub> O to 1 litre (approximate pH of 8.5)
High Ionic Strength 4 % Gel Mix	8.0ml 5x Tris-glycine, 5.33ml 30 % (w/v) Acrylamide, 1ml 2 % (w/v) N-M-methylene-bisacrylamide, 2.0ml 50 % (v/v) Glycerol, 23.7ml H <sub>2</sub> O. Filtered and used

within 24 hrs of preparation. 200µl 15% ammonium persulphate and 34µl Temed added before use

High Ionic strength                      1x Tris-glycine

Electrophoresis Buffer

5x Gel Shift Loading                      8ml 5x Tris-glycine, 2ml glycerol, 0.01g Bromophenol  
Buffer    Blue

## **2.4        METHODS**

### **2.4.1     MICROBIOLOGICAL TECHNIQUES**

#### **2.4.1.1   Growth of *E.coli* Bacterial Cultures**

Liquid culture of *E.coli* was either in Luria broth, 2xTY broth (HB101) or minimal medium (NM522), containing ampicillin at a final concentration of 100µg/ml if required. 5mls of the appropriate broth in 2oz glass bottles was inoculated with a single colony using a sterile inoculating loop or a sterile toothpick. Cultures greater than 10mls were grown in conical flasks with a capacity of 5-10x that of the culture volume. Cells were shaken at 37°C for an appropriate length of time.

#### **2.4.1.2   Storage of *E.coli* Bacterial Cultures**

For short-term storage (4-6 weeks), bacteria were stored as 5ml overnight liquid cultures or streaked onto agar plates, containing ampicillin at a final concentration of 100µg/ml if required.

For long-term storage, 500µl from a fresh overnight culture of bacteria was mixed with 500µl of 100 % glycerol (autoclaved), and stored in a sterile eppendorf at -70°C. Upon recovery, vials were thawed quickly at 37°C and the culture streaked out onto agar plates, containing antibiotic if required. NM522 was



streaked onto minimal agar to maintain the F' plasmid. After incubation at 37°C, a single colony was used to propagate a fresh bacterial culture.

#### **2.4.1.3 Preparation and Transformation of Competent Cells**

Several methods were employed to prepare competent cells depending on the transformation procedure to be used. 5 to 10ng of supercoiled plasmid DNA, or half a ligation reaction, were used to transform competent cells.

**a) CaCl<sub>2</sub> Method** - 100ml of Luria broth supplemented with 1ml 20 % (w/v) glucose was inoculated with 1ml of a fresh overnight culture of the appropriate bacterial strain and grown at 37°C with shaking until an OD<sub>650nm</sub> of 0.45-0.55 was reached for HB101, or 0.3-0.4 for NM522. The culture was placed on ice for 10 minutes to arrest growth before pelleting the cells by centrifugation at 2000rpm for 10 minutes at 4°C. The cells were resuspended in 50ml of 50mM sterile, ice-cold CaCl<sub>2</sub> and incubated on ice for 15 minutes before re-pelleting and resuspending in 2.5ml of 50mM sterile, ice-cold CaCl<sub>2</sub>. At this stage, the cells were divided into 100µl aliquots which could be used for transformation immediately, or snap frozen in liquid nitrogen (or dry ice in ethanol) and stored indefinitely at -70°C for later use.

For transformation, 100µl of the cell suspension was thawed at room temperature, added to a ligation mixture and incubated on ice for 30 minutes. The cells were then heat-shocked at 42°C for 2 minutes and returned to ice for 1 minute. 0.9ml of Luria broth was added to each transformation mixture and the cells incubated at 37°C for 1 hour. 200µl from each transformation were plated onto appropriate agar plates, containing ampicillin to a final concentration of 50µg/ml if required, and incubated at 37°C overnight.

**b) DMSO Method** - 50ml of Luria broth supplemented with 500µl of 20 % (w/v) glucose solution was inoculated with 500µl of a fresh overnight culture of the relevant bacterial strain and grown at 37°C until an OD<sub>650nm</sub> of 0.45-0.55 was reached for HB101, or 0.3-0.4 for NM522. The culture was placed on ice for 10 minutes to arrest growth and the cells pelleted by centrifugation at 2000rpm for 10

minutes at 4°C. The cells were resuspended in 1/10th volume (5ml) of Transformation and Storage Solution, or TSS (Luria broth containing 10 % (w/v) PEG (molecular weight 3000), 5 % (v/v) DMSO, 10mM MgCl<sub>2</sub> and 10mM MgSO<sub>4</sub>). The cells were incubated on ice for 10 minutes after which they were divided into 100µl aliquots for immediate use, or snap cooled in liquid nitrogen (without the addition of glycerol) for indefinite storage at -70°C.

For transformation, 100µl of the cell suspension was thawed at room temperature, added to a ligation mixture and incubated on ice for 1 hour. 0.9ml of TSS and 20µl of 20 % (w/v) glucose solution were added to each transformation mixture and the cells incubated at 37°C for 1 hour. 200µl from each transformation were plated onto appropriate agar plates, containing ampicillin to a final concentration of 50µg/ml if required, and incubated at 37°C overnight.

**c) Electroporation Method** - 1 litre of Luria broth was inoculated with 10ml of a fresh overnight culture of NM522 and grown at 37°C until an OD<sub>260nm</sub> of 0.5-1.0 was reached. The culture was placed on ice for 10 minutes before pelleting the cells by centrifugation at 2000rpm for 10 minutes at 4°C. The cell pellet was resuspended in 1 litre of ice-cold sterile water and the centrifugation repeated. The procedure of resuspending and pelleting the bacterial cells was repeated using 500ml and 100ml of sterile water, and then 20ml of sterile 10 % (v/v) glycerol, before finally resuspending in 2-3ml of 10 % glycerol. At this stage, the cells were divided into 40µl aliquots which could be used for transformation immediately or snap frozen in liquid nitrogen and stored at -70°C indefinitely.

For transformation, 40µl of the cell suspension was thawed at room temperature and added to 1µl or 2µl of DNA, which was in a low ionic strength buffer such as TE (ligation reaction mixtures were phenol/chloroform extracted, ethanol precipitated and resuspended in 2µl TE). The cell suspension and DNA were mixed well and stored on ice for 1 minute. The Gene Pulser apparatus (BioRad) was set at 25µFD and 2.5kV, and the Pulse Controller at 200Ω. The transformation mixture was transferred to a cold, 0.2cm electroporation cuvette and shaken to the bottom before placing in a safety chamber slide. The slide was pushed into the

chamber until the cuvette contacted the conductors at the base of the chamber. Electroporation with the above settings produces a pulse with a time constant of 4.5 to 5.0 milliseconds and a field strength of 12.5kV/cm. Immediately following electroporation the cuvette was removed from the chamber for the addition of 1ml of SOC medium (2 % Bacto Tryptone, 0.5 % Bacto Yeast Extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl<sub>2</sub>, 10mM MgSO<sub>4</sub>, 20mM glucose). The cell suspension was transferred to an eppendorf and incubated at 37°C for 1 hour after which the cells were spun down briefly in a microfuge, resuspended in 200µl of Luria broth and plated onto appropriate agar plates, containing ampicillin to a final concentration of 50µg/ml if required, and incubated overnight at 37°C. To select for recombinants using the blue/white screen 40µl of X-gal (25mg/ml in dimethylformamide) and 40µl of IPTG (25mg/ml in sterile H<sub>2</sub>O) were added to each agar plate.

#### **2.4.1.4 Small Scale Preparation of Plasmid DNA**

**a) Alkaline Lysis Method** - A single plasmid carrying bacterial colony was used to inoculate 5ml of Luria broth, supplemented with ampicillin to a final concentration of 50µg/ml, which was grown, with shaking, at 37°C overnight. A 1.5ml aliquot of this culture was pelleted by centrifugation in an Eppendorf microfuge (12Krpm, 1 minute) and the supernatant discarded. Cells were resuspended in 100µl of Solution A (25mM Tris-HCl pH 8.0, 10mM EDTA, 50mM glucose), vortexed and incubated at room temperature for 5 minutes. 200µl of lysis buffer (0.2M NaOH, 1 % SDS) was added and after gentle mixing by inversion, incubated on ice for 10 minutes. Following cell lysis, 150µl of 3M sodium acetate (pH 5.2) was added and mixed by gentle inversion to precipitate chromosomal DNA, SDS and proteins. Following further incubation on ice for 10 minutes the precipitate was centrifuged in a microfuge for 5 minutes and the supernatant transferred to a fresh Eppendorf, taking care not to disturb the precipitate. The sample was then phenol/chloroform/iso-amyl alcohol extracted (2.4.2.1) and the plasmid DNA precipitated with ethanol (2.4.2.2). The DNA pellet was allowed to air-dry before resuspending in 20µl TE buffer. Aliquots of 2µl were used in restriction analysis

(2.4.2.3) with RNaseA added to a final concentration of 0.1µg/µl in each restriction digest reaction.

**b) Rapid Boiling Method** - DNA was prepared as for the Alkaline Lysis method. However, instead of resuspending the bacterial pellet in 100µl of Solution A, it was resuspended in 350µl of Boiling Buffer (10mM Tris-HCl pH 8.0, 50mM EDTA pH 8.0, 8 % (w/v) sucrose, 0.5 % (v/v) Triton X-100). 25µl of freshly prepared lysozyme (10mg/ml in 10mM Tris-HCl pH 8.0) was added and the suspension vortexed. The sample was then placed in a boiling water bath for 40 seconds and immediately centrifuged in a microfuge for 10 minutes. The pellet was removed using a sterile toothpick. The supernatant was then phenol/chloroform/iso-amyl alcohol extracted (2.4.2.1) and ethanol precipitated (2.4.2.2) to obtain plasmid DNA. The DNA was washed with 70 % ethanol and allowed to air dry before resuspending in 50µl TE buffer. 5-10µl was used for restriction analysis (2.4.2.3) with RNaseA added to a final concentration of 0.1µg/µl in each digest.

#### **2.4.1.5 Large Scale Preparation of Plasmid DNA**

**a) CsCl<sub>2</sub> Method** - 2.5ml from an overnight culture of bacteria containing the required plasmid was used to inoculate 250ml Luria broth, containing ampicillin at a final concentration of 100µg/ml. This was grown overnight with shaking at 37°C. The cells were pelleted in a Sorval RC-5B centrifuge (using a Sorval HS-4 GSA rotor) at 3Krpm for 10 minutes at 4°C. The supernatant was discarded and the cell pellet resuspended thoroughly in 5ml of ice cold Solution A (25mM Tris-HCl pH 8.0, 10mM EDTA, 50mM glucose). 100µl of lysozyme (5mg/ml in 10mM Tris-HCl pH 8.0) was added to the cell suspension, vortexed, and incubated on ice for 5 minutes. Cells were lysed by adding 10ml of cold (but not precipitated) lysis buffer (0.2M NaOH, 1 % SDS), vortexing and incubating on ice for 5 minutes, after which 7.5ml of ice cold 3M sodium acetate (pH 5.2) was added to precipitate chromosomal DNA, SDS and proteins. This was mixed carefully and stored on ice for 10 minutes. The sample was then transferred to a 30ml corex tube and the precipitate spun down

at 10Krpm (using a Sorval SS-34 rotor) for 30 minutes at 4°C. The supernatant (25ml) was transferred to a 50ml Falcon tube, to which 15ml (0.6 volumes) of ice cold isopropanol was added and left to precipitate at -20°C for 1 hour or overnight. The DNA solution was transferred to 2 X 30ml corex tubes and pelleted by centrifugation at 10Krpm (Sorval SS-34 rotor) for 20 minutes at 4°C. The pellet was washed with 70 % ethanol, allowed to dry, resuspended in 9.5ml TE buffer plus 10µg/ml RNaseA and incubated at 37°C for 30 minutes. 10g of caesium chloride and 200µl of ethidium bromide (15mg/ml) were added and the solution centrifuged at 38Krpm (using a Beckman Ti50 rotor) for 40-48 hours at 20°C. DNA bands were visualised by side illumination with UV light. A 19 or 21 gauge needle was used to puncture the top of the tube to release pressure. A second needle was used to remove the lower band containing supercoiled plasmid DNA. The ethidium bromide was extracted several times by the addition of equal volumes of butan-2-ol saturated with water, until the plasmid DNA solution became clear. The volume of the DNA solution was made up to 5ml with TE buffer and precipitated by adding 10mls of ice cold absolute ethanol and storing at -20°C for 1 hour or overnight. The sample was then centrifuged at 10Krpm (Sorval SS-34 rotor) for 10 minutes at 4°C, the DNA pellet washed with 70 % ethanol and resuspended in 200-500µl TE buffer. Finally, the DNA concentration was determined (2.4.1.6). DNA to be used for P-element mediated germ line transformation was further phenol/chloroform/iso-amyl alcohol extracted (2.4.2.1), ethanol precipitated (2.4.2.2) and resuspended in injection buffer at the appropriate concentration (2.4.3.1).

**b) QIAGEN Plasmid Maxi Kit** - Qiagen kits were obtained from Hybaid. Up to 500µg of double stranded supercoiled DNA was isolated from *E.coli* cultures following the protocol given in the QIAGEN plasmid handbook, which provided suitable template for double stranded plasmid dideoxynucleotide sequencing (2.4.2.9).



#### **2.4.1.6 Estimation of Plasmid DNA Yields**

A measurement of the optical density (OD) at 260nm was used to calculate the concentration of DNA in a given solution. An OD<sub>260</sub> of 1.0 is equivalent to 50µg/ml for dsDNA and 40µg/ml for ssDNA or RNA. Nucleic acid was considered to be free of contaminating protein if the OD<sub>260</sub>/OD<sub>280</sub> was 1.8-2.0 since a ratio of 1.8 indicates purity for a DNA sample and a ratio of 2.0 indicates purity for an RNA sample. Estimates of DNA concentrations were also obtained by visual comparison of DNA samples to a known quantity of material in ethidium bromide stained agarose gels (2.4.2.4).

### **2.4.2 MANIPULATION AND DETECTION OF NUCLEIC ACIDS**

#### **2.4.2.1 Phenol and Chloroform Extraction of Proteins from Nucleic Acid**

Distilled phenol saturated with water (Raithburn Chemicals) was equilibrated to pH 8.0 by mixing the phenol with an equal volume of 1M Tris-HCl pH 8.0. The phases were allowed to separate and the upper aqueous phase was discarded. This process was repeated until phenol reached pH 8.0, after which the Tris layer was replaced with TE buffer. DNA solutions were phenol extracted by mixing thoroughly with an equal volume of phenol until an emulsion formed. The phases were then separated by centrifugation in a microfuge (12Krpm, 2-5 minutes), and the aqueous phase transferred to a fresh Eppendorf, taking care not to disturb the proteins located at the interface. Instead of, or following phenol extraction, DNA was extracted with phenol/chloroform/iso-amyl alcohol (25:24:1). The aqueous phase was finally extracted with an equal volume of chloroform/iso-amyl alcohol (24:1) to remove any traces of phenol, and the aqueous layer transferred to a fresh tube

#### **2.4.2.2 Precipitation of DNA**

DNA was precipitated from solution by adding 1/10th volume of 3M sodium acetate or potassium acetate (pH 4.8-5.2) and twice the volume of ice cold absolute ethanol or 0.6 volumes of ice cold propan-2-ol. Samples were then stored at

-20°C for a minimum of 1 hour, or -70°C for 20 minutes, or incubated in a dry ice/ethanol bath for 10 minutes. The DNA was recovered by centrifugation in a Sorval centrifuge (10Krpm, 10 minutes, 4°C) or a microfuge (12Krpm, 10 minutes, room temperature). The supernatant was discarded and the DNA pellet washed with 70 % ice cold ethanol and spun again for 2 minutes. The supernatant was discarded and the DNA pellet allowed to dry until no visible traces of ethanol remained (2-20 minutes). The DNA was re-dissolved in TE buffer of neutral pH.

#### **2.4.2.3 Restriction Endonuclease Digestion of DNA**

Reaction conditions recommended by the manufacturers were used with the correct buffer supplied. 0.5-5.0µg of DNA was cut in a 10-40µl reaction volume containing the appropriate volume of 10x buffer required, for 1-2 hours at 37°C. For double digests involving enzymes with different recommended buffers, the buffer which gave the most efficient digestion was used. When RNA was expected to be present in the DNA sample (as with the small scale preparations of plasmid DNA; 2.4.1.4) RNaseA was added to the reaction to a final concentration of 0.1µg/µl.

#### **2.4.2.4 Agarose Gel Electrophoresis of DNA**

Agarose gels (0.7-1.2 %) were made and run in 1x TBE gel buffer. Ethidium bromide was added at a final concentration of 0.5µg/ml to the gel. DNA samples were mixed with 1/10th volume of 10 X DNA gel loading buffer. NBL and Biorad mini gel, and NBL midi gel apparatus was used. Mini-gels were run at 50-100V for 1-2 hours and midi-gels were run at 100-150V for 1-4 hours, or 25-50V for 16-20 hours. After electrophoresis, DNA was visualised on a chromatovue C-70G UV viewing system (254nm short wavelength, 265nm long wavelength) and photographed.

#### 2.4.2.5 Molecular Weight Markers

The sizes of DNA restriction fragments were determined by comparing the distance migrated in the gel to that of the distance migrated by DNA fragments of known molecular weight. *Hind*III digested  $\lambda$ DNA (NBL) and a set of DNA fragments differing by approximately 1 kb in size (1 kb ladder from Gibco BRL) were used as standard molecular weight markers.

#### 2.4.2.6 Recovery of DNA from Agarose Gels

a) **GeneClean** - GeneClean kits were supplied by BIO 101 Inc., and allow the purification of DNA fragments by means of DNA binding affinity for a silica matrix. The efficiency of recovery for DNA fragments of over 500 bp is 60-70 %. Following gel electrophoresis, the desired DNA fragment was visualised and cut out of the agarose gel. The agarose was weighed such that 2.5 volumes of 6M NaI solution and 0.5 volumes of TBE modifier solution could be added. The agarose was dissolved by heating to 55°C for 5 minutes with occasional mixing, and then cooled on ice for 5 minutes. 2-5 $\mu$ l of glassmilk (a silica matrix suspended in water) was added, and stored for 5 minutes on ice (with occasional careful mixing) to allow the DNA to bind to the silica matrix. The glassmilk/DNA was pelleted by centrifugation for 30 seconds in a microfuge and washed three times with an appropriate volume (100 X glassmilk volume) of NEW wash (NaCl/ethanol/water mix). After the final centrifugation, all of the NEW wash was removed and the DNA eluted from the glassmilk in 5-10 $\mu$ l of TE buffer by incubation at 55°C for 10-15 minutes. The glassmilk was removed by centrifugation for 1 minute in a microfuge. The supernatant, containing the DNA, was transferred to a fresh Eppendorf and the centrifugation repeated to remove all traces of the silicon matrix before using the DNA in ligation reactions, as glassmilk could inhibit this process. The remaining DNA was stored at -20°C.



**b) Qiaex DNA Gel Extraction** - Qiaex kits were supplied by Hybaid and also employ a resin that binds DNA to purify DNA fragments. DNA fragments from as little as 50 bp can be purified with an efficiency of recovery similar to that of GeneClean. The desired DNA fragment was excised from an agarose gel and weighed before adding 300µl of buffer QX1 per 100mg of gel, 0.1M Mannitol and 10µl of the resin suspension. Incubation at 50°C for 10 minutes allowed the gel to melt and the DNA fragment to bind to the resin. The resin was spun down for 30 seconds in a microfuge and washed twice with 500µl of buffer QX2, and twice with 500µl of buffer QX3, pelleting by centrifugation between washes. Following the last wash, the supernatant was discarded and the pellet allowed to air-dry. The DNA was eluted in 10-20µl of TE buffer by incubation at 50°C for 10-15 minutes and used directly in ligation reactions, or stored at -20°C for future use.

#### **2.4.2.7 Ligation of DNA Molecules**

50-100ng of linearised vector (cut to completion with the appropriate restriction enzyme) was incubated with a 0.5x, 1x and 2x molar ratio of insert DNA in 1x ligation buffer (10mM Tris-HCl pH 7.2, 1mM EDTA, 10mM MgCl<sub>2</sub>, 10mM DTT, 1mM ATP) with 1-10 units of T4 DNA ligase. The reaction was made up to a final volume of 10µl and left overnight at 16°C.

**a) Vector dephosphorylation** - If the vector was linearised using a single enzyme, self ligation was prevented by removing the terminal 5'-phosphate groups with the addition of 0.01 units of calf intestinal alkaline phosphatase (CIAP) for 15 minutes at the end of the restriction digest reaction. Prior to ligation, CIAP was removed by phenol/chloroform/iso-amyl alcohol extraction (2.4.2.1) and the DNA ethanol precipitated (2.4.2.2).

**b) Creation of blunt ends** - DNA fragments with 5' overhangs were blunt-ended by end-filling the 3' recessed termini. Restriction digests (1-2µg DNA) were carried out in a volume of 20µl and following digestion 3µl 2mM dNTPs, 1µl restriction buffer used in the digest and 2µl Klenow (1u/µl) were added, the volume

made up to 30µl with TE buffer, and the reaction allowed to proceed at room temperature for 30 minutes. The reaction was stopped by the addition of 1µl 0.5M EDTA pH 8.0, made up to 100µl with TE buffer, phenol/chloroform/iso-amyl alcohol extracted (2.4.2.1) and ethanol precipitated (2.4.2.2). DNA fragments with 3' protruding termini were blunt ended using the 3' → 5' exonuclease activity of bacteriophage T4 DNA polymerase. Following digestion, samples were phenol/chloroform/iso-amyl alcohol extracted, ethanol precipitated and resuspended in 10µl TE buffer. To this was added 2µl 10x T4 polymerase buffer (0.33M Tris acetate pH 7.9, 0.66M potassium acetate, 0.1M magnesium acetate, 5mM dithiothreitol, 2.5µg/µl BSA), 1µl 2mM dNTPs, 2.5 units T4 DNA polymerase and TE buffer to 20µl. Samples were incubated for 30 minutes at room temperature and reactions stopped by adding 1µl 0.5M EDTA pH 8.0. Samples were deproteinised by phenol/chloroform/iso-amyl alcohol extraction and ethanol precipitation or immediately loaded onto agarose gels for electrophoresis (2.4.2.4) and gel purification (2.4.2.6).

**c) Addition of linkers to insert DNA** - A 50-fold molar excess of phosphorylated *EcoRI* linkers were incubated with 200ng of insert DNA (blunt-ended) in 1x ligation buffer with 2.5 units of T4 DNA ligase. BSA was included in each reaction at a final concentration 0.1µg/µl. Each reaction made up to a final volume of 10µl and incubated at 16°C overnight. Ligase was inactivated by incubating the reactions at 70°C for 10 minutes. Tubes were cooled on ice before digesting samples with *EcoRI* (1 unit per pmole linkers used in the ligation reaction). Following incubation at 37°C for 2 hours, samples were deproteinised by phenol/chloroform/iso-amyl alcohol extraction (2.4.2.1) and the insert DNA (with linkers attached) ethanol precipitated (2.4.2.2). Samples were resuspended in 2µl TE buffer, and cloned into the appropriate vector.

When sufficient quantities of DNA were available prior to ligation reactions, both vector and insert DNA concentrations were estimated by visual comparison to a known quantity of DNA on ethidium bromide stained agarose mini-gels (2.4.2.4). Ligation products were transformed into *E.coli* (2.4.1.3).

#### **2.4.2.8 Polymerase Chain Reaction**

PCR reactions were carried out using a Techne Programmable Dryblock. Primers were designed that shared homology to 20 nucleotides in the template and had a 6 bp restriction enzyme site preceded by a G at their 5' end. 100pg of template DNA was mixed with 5µl of 10x amplification buffer (100mM Tris-HCl pH 8.3, 500mM KCl, 15mM MgCl<sub>2</sub>, 0.1 % (w/v) gelatin, 0.1 % (v/v) tween-20, 0.1 % (v/v) NP40), 2µl of 10mM dNTPs, 1-2µg of each primer and 2.5 units of Taq polymerase (Promega). The reaction mixture was made up to 50µl with sterile water and submerged under 50µl of mineral oil (Sigma). Three denaturation cycles of 94°C for 1 minute, annealing at 55°C for 1 minute and polymerisation at 72°C for one minute were carried out. A further 20 cycles of 93°C for 30 seconds, 55°C for 30 seconds and 72°C for 30 seconds were done. The amplification process was completed with a cycle of 93°C for 30 seconds, 55°C for 30 seconds and 72°C for 5 minutes. PCR products were checked by agarose gel electrophoresis (2.4.2.4) and excess primers removed by isopropanol precipitation (one volume of isopropanol and 0.1 volumes of sodium acetate pH 4.8, mixed and centrifuged for 10 minutes in a microfuge). The DNA pellet was washed and dried as in ethanol precipitation procedure (2.4.2.2).

#### **2.4.2.9 Dideoxynucleotide Sequencing of Double Stranded Plasmid DNA**

DNA sequencing was carried out using the Sequenase<sup>TM</sup> Version 2.0 kit (United States Biochemicals).

i) Template Preparation - 5µg of plasmid DNA (prepared using the QIAGEN Plasmid Maxi Kit - 2.4.1.5b) in 18µl of TE buffer was denatured by adding 2µl of 2M NaOH and incubating at room temperature for 15 minutes. The DNA was neutralised by adding 5.3µl of 7.5M ammonium acetate, and ethanol precipitated by the addition of 100µl chilled absolute ethanol and placing the mixture in a dry ice/ethanol bath for 5 minutes. The DNA was pelleted by spinning in a microfuge for 5 minutes, washed in 80 % ethanol and allowed to air-dry. The DNA was resuspended in 7µl sterile distilled water for use in the sequencing reaction:

ii) Annealing Reaction - Primer was annealed to the single stranded template by adding 2µl of 5x sequenase reaction buffer (200mM Tris-HCl pH 7.5, 100mM MgCl<sub>2</sub>, 250mM NaCl), 1µl of primer (20ng/µl) and heating at 37°C for 10 minutes, before storing on ice.

iii) Labelling Reaction - Extension from the annealed primer was done by adding 1µl 0.1M DTT, 2µl dGTP label mix (a 1 in 4 dilution of 7.5µM dGTP, dCTP and dTTP), 0.5µl α-<sup>35</sup>S[dATP] (400Ci/mmole) and 2µl Sequenase<sup>TM</sup> (a 1 in 8 dilution of Sequenase at 13 u/µl in 10mMTris-HCl pH 7.5, 5mM DTT, 0.5mg/ml BSA). The reaction components were mixed by briefly spinning in a microfuge and incubated at 37°C for 5 minutes.

iv) Termination Reaction - For each extension or labelling reaction, four tubes were preheated to 37°C, containing 2.5µl of one of the four termination mixes:

ddTTP mix - 80µM dNTPs, 8µM ddTTP, 50mM NaCl

ddCTP mix - " " ddCTP, "

ddGTP mix - " " ddGTP, "

ddATP mix - " " ddATP, "

3.5µl of the labelling reaction was added to each tube and the termination reaction allowed to proceed at 37°C for 5 minutes. The reactions were stopped on the addition of 4µl of stop solution (95 % Formamide, 20mM EDTA, 0.05 % Bromophenol Blue, 0.05 % Xylene Cyanol FF).

#### **2.4.2.10 Polyacrylamide Gel Electrophoresis of DNA Sequencing Reactions**

Sequencing gels (45cm x 18cm x 0.2mm, or 60cm x 18cm x 0.2mm) were prepared with 6 % (w/v) acrylamide (19:1 acrylamide:bis-acrylamide) in TBE buffer containing 7.7M urea. Polymerisation of the acrylamide was achieved by adding ammonium persulphate and Temed (2.3.1) immediately prior to pouring the gel. Sequencing reactions were denatured by boiling for 3 minutes, loaded onto the gel and electrophoresed in 1x TBE buffer at a constant power of 35 watts for 2-3 hours (for the 45cm x 18cm x 0.2mm gel) or 45 watts for 3-6 hours (for the 60cm x 18cm x 0.2mm gel). Following electrophoresis, gels were fixed in 10 % (v/v) methanol, 10 % (v/v) acetic acid for 15-20 minutes, transferred onto blotting paper

(Ford Goldmedal Blotting Paper), dried under vacuum at 80°C for 1-2 hours (Biorad gel drier, model 583) and autoradiographed at room temperature overnight.

### **2.4.3 MANIPULATIONS OF *DROSOPHILA* STOCKS**

#### **2.4.3.1 Establishment of Transgenic Lines by P-Element Transformation**

P-element mediated germ line transformation (Rubin and Spradling, 1982) of W<sup>K</sup> *Drosophila* embryos was performed with DNA prepared as described in section 2.4.1.5a. The transformation vector containing the DNA sequence to be tested was mixed with the helper plasmid phs70Δ2-3wc at a ratio of 4:1. The DNAs were phenol/chloroform/iso-amyl alcohol extracted (2.4.2.1), ethanol precipitated (2.4.2.2) and resuspended in injection buffer (5mM KCl, 0.1mM NaPO<sub>4</sub> pH6.8) to give a final concentration of 400μg/ml of transformation construct and 100μg/ml of helper plasmid. The DNA was centrifuged for 2 minutes in a microfuge to spin down any debris, taken up into an injection needle and stored under Kel-F halocarbon oil to prevent drying. Injection needles were pulled from 25μl Drummond glass microcapillaries using a vertical pipette puller (Model 700C, David Kopf Instruments, Tujunga, California), mounted in the micromanipulator (Leitz) and a sharp point produced by fracturing the tip against the edge of a cover slip or razor blade (but you have to be good to do it with the razor blade) adhered to a microscope slide with double sided tape.

Eggs were collected every 30 minutes from 3-7 day old W<sup>K</sup> flies kept in an egg collecting chamber with a sliding food drawer to allow easy removal of newly laid eggs. Eggs were dechorionated manually by gently rolling across double sided tape (Scotch brand 3M No.666) mounted on a microscope slide. The dechorionated eggs were lined up on a narrow strip of tape, close to the edge of the microscope slide, with their posterior ends pointing in the same direction, and slightly dessicated by placing the microscope slide in a petri dish containing silica gel dessicant for 3-6 minutes (depending on the humidity and temperature of the room, and how long it took to dechorionate them). The embryos were then covered with Kel-F halocarbon oil and mounted on the stage of the micromanipulator. DNA

was gently injected into the posterior pole of preblastoderm eggs prior to pole cell formation, or stage 1 and 2 embryos (Bownes, 1975). Embryos at later stages of development were destroyed by tearing the vitelline membrane. Injected eggs were incubated in a humid chamber at 18°C for 36-60 hours, and larvae collected as they emerged. Larvae were placed in vials containing yeast solution and fly food, and incubated at 25°C. When any adults eclosed, they were individually mated with 3-5 W<sup>K</sup> virgins, and the progeny examined for a red eye phenotype. Individual transformants were crossed with a single W<sup>K</sup> virgin and homozygous lines established by sibling matings with the progeny from this cross. Stably transformed lines were characterised by Southern blot analysis.

#### **2.4.3.2 Preparation of Genomic DNA from Adult Flies**

A variation on a rapid method for DNA extraction from *Drosophila* was used (Marcus, 1985). Twenty flies were etherised before freezing in liquid nitrogen, at which stage they could be stored at -70°C until use. Flies were homogenised in 350µl of DNA Extraction Buffer I (0.5M Tris-HCl pH 8.0, 0.15M NaCl, 15mM EDTA) in a 1.5ml Eppendorf tube, using a hand held motorised homogeniser. After this stage, care was taken to prevent shearing of genomic DNA. 350µl of DNA Extraction Buffer II (0.5M Tris-HCl pH 8.0, 0.15M NaCl, 15mM EDTA, 0.04 % (w/v) SDS) was added, and Proteinase K to a final concentration of 50µg/ml. This was mixed by gentle pipetting using a 1ml Gilson pipette with a cut-off blue tip. The homogenate was incubated at 65°C for 1 hour, phenol/chloroform/iso-amyl alcohol extracted twice and chloroform/iso-amyl alcohol extracted once (2.4.2.1), shaking by hand to mix. Genomic DNA was precipitated by adding 1/10th volume of 3M sodium acetate (pH 5.2) and 0.6 volumes of chilled isopropanol, and left for at least 3-4 hours at -20°C. DNA was pelleted in a microfuge for 10 minutes, washed with 70 % chilled ethanol, allowed to air-dry and resuspended in 24µl TE buffer either by gentle pipetting or by leaving at 4°C overnight. Genomic DNA was digested for 4-8 hours with 10-20 units of the chosen restriction enzyme, in appropriate buffer conditions (2.4.2.3), with RNaseA added to a final concentration of 0.1µg/µl.

### 2.4.3.3 Southern Blotting (Southern, 1975)

*Drosophila* genomic DNA samples were run in 0.8 % (w/v) agarose gels (2.4.2.4) at 30-40V overnight. Gels were photographed before being gently shaken for 30 minutes in Denaturation Buffer (1.5M NaCl, 0.5M NaOH), rinsed in water, gently shaken for 30 minutes in Neutralisation Buffer (1.5M NaCl, 0.5M Tris-HCl pH 7.2, 1mM EDTA) and washed in 2x SSC.

To set up the blot, a blotting paper wick was placed on a platform supported above a reservoir of 20x SSC. The wick was saturated with 20x SSC, the gel inverted and placed on top. Strips of parafilm were placed around the edges of the gel, between the wick and the gel, to insulate the wick from the rest of the set up. A sheet of Hybond-N membrane filter (Amersham) of the same dimensions as the gel, and wetted with 2x SSC, was placed on top of the inverted gel taking care to avoid trapping air bubbles. 3-5 pieces of blotting paper cut to the same dimensions as the gel were soaked in 2x SSC and placed on top of the Hybond-N membrane. Further layers of dry blotting paper, also of the same dimensions as the gel, were placed on top of the wet blotting paper to a thickness of 3-5cm. Finally, a stack of paper towels were added to a thickness of 5cm. A sheet of perspex was placed over this and a weight placed on top, secured by a clamp if necessary. The capillary blot transfer was left to proceed overnight (16-24 hours) after which the membrane filter was removed, rinsed briefly in 2x SSC, allowed to air dry and wrapped in clingfilm. The DNA was fixed to the filter by UV irradiation for 10 minutes (ChromatoVue C-70G, UV viewing system - 254nm short  $\lambda$  and 365nm long  $\lambda$ ) with the DNA side facing the UV light source. Membranes could be sealed in a plastic bag for storage or used immediately for hybridisation.

If Hybond-N<sup>+</sup> membrane was used in the blot, the transfer of DNA from the gel onto the membrane was achieved using an alkali transfer procedure. The gel was soaked for 30 minutes with gentle shaking in 0.25M HCl (43.2ml conc. HCl in 2 litres H<sub>2</sub>O), rinsed in distilled water for 5 minutes and placed in 0.4M NaOH for 5 minutes before use in the blot. Instead of using 20x SSC, the transfer was done with 0.4M NaOH. The DNA is denatured in the transfer process such that

gel pre-treatment steps are less involved and UV fixing of the DNA to the filter is not required.

#### **2.4.3.4 Radioactive Labelling of DNA by Random Priming (Feinberg and Vogelstein, 1983, 1984)**

200ng of the DNA fragment, in a total volume of 35 $\mu$ l, to be labelled was denatured by boiling for 5 minutes and snap cooled on ice for 5 minutes. To this was added 1 $\mu$ l BSA (20mg/ml), 10 $\mu$ l OLB (detailed in section 2.3.2), 3 $\mu$ l (30 $\mu$ Ci)  $\alpha$ -<sup>32</sup>P[dCTP] and 1 $\mu$ l of *E.coli* DNA polymerase I Klenow fragment (5u/ $\mu$ l). The reaction was incubated at room temperature for 4-16 hours. Radiolabelled DNA fragments were separated from unincorporated nucleotides by passing through a 1ml Sephadex G-50 NICK<sup>TM</sup> Column (Pharmacia) that was rinsed and equilibrated with 3ml TE buffer. The labelled DNA was then added to the Column followed by 400 $\mu$ l TE buffer. A further 400 $\mu$ l TE buffer was then used to elute the labelled DNA sample.

#### **2.4.3.5 Hybridisation of Southern Blot Filters**

Hybond-N and Hybind-N+ membranes used for Southern blots were wrapped in Hybaid membranes and incubated with 10-20mls of pre-hybridisation buffer (0.5M Na<sub>2</sub>HPO<sub>4</sub> pH 7.0, 7 % (v/v) SDS, 50 % (v/v) Formamide) in Hybaid oven bottles at 42°C for at least 2 hours. The radiolabelled DNA was denatured by incubating at 95°C for 10 minutes and snap-cooled on ice for 5 minutes before adding to the Hybaid bottle containing the pre-hybridised filter and solution. Hybridisation was allowed to proceed at 42°C overnight. Following hybridisation the solution containing the probe was discarded and the filters washed in solutions with gradually increasing stringency conditions. After each wash under a particular stringency the filter was monitored to check the level of the radioactive signal, before proceeding to a higher stringency wash. Filters were rinsed with 25ml of 2x SSC at 42°C for 15 minutes, washed twice with 25ml Low Stringency Buffer (2x SSC, 0.1 % SDS) at 42°C for 15 minutes, twice with 25ml Medium Stringency Buffer (1x SSC, 0.1 % SDS) at 42°C for 15 minutes, twice with 25ml High



Stringency Buffer (0.1x SSC, 0.1 % SDS) at 42°C for 15 minutes and, if necessary, once in 25ml High Stringency Buffer at 65°C for 15 minutes. Filters were finally sealed in a polythene bag for autoradiography.

#### **2.4.3.6 Histochemical Staining of Transformed *Drosophila***

Flies were dissected in Ringer's Solution (55mM NaCl, 40mM KCl, 7mM MgSO<sub>4</sub>, 5mM CaCl<sub>2</sub>, 10mM tricine, 20mM glucose, 50mM sucrose) and stained for  $\beta$ -galactosidase activity overnight at room temperature in 100 $\mu$ l of X-gal Staining Buffer (detailed in section 2.3.4).  $\beta$ -galactosidase activity represents *lacZ* reporter gene expression, indicated by blue colouration appearing over several hours.

#### **2.4.3.7 Heat Shock Treatment of *Drosophila***

Flies transformed with the pERI transformation vector were heat-shocked by incubating at 37°C for 45 minutes and allowed to recover for 1 hour at room temperature prior to histochemical staining.

#### **2.4.3.8 20-hydroxyecdysone Treatment of Transformed *Drosophila***

Males were etherised and injected in the abdomen with 0.1 $\mu$ l of 10<sup>-2</sup>M 20-hydroxyecdysone (2.3.5) and maintained in vials in a humid atmosphere either at 18°C overnight or 25°C for 6 hours, prior to analysis. For treatment of females, flies were starved on 3 % (w/v) agar for 48 hours prior to injection with 20-hydroxyecdysone, and allowed to recover at 18°C overnight.

#### **2.4.3.9 Extraction of Haemolymph Samples from *Drosophila***

Flies were lightly etherised and attached by their wings to a microscope slide using double sided adhesive tape. The haemolymph was collected with a glass needle (prepared as described in 2.4.3.1) through the anterior ventral surface of the thorax and stored in 60 $\mu$ l 2x polypeptide sample buffer (2.3.3). This was frozen in liquid nitrogen before storing at -20°C, until ready for SDS-PAGE analysis.

#### 2.4.3.10 SDS-Polyacrylamide Gel Electrophoresis of Proteins (Laemmli, 1970)

A 2mm 10 % polyacrylamide separating gel (2.3.3) was poured between glass plates until three quarters of the available volume was occupied. Overlay buffer (0.1 % SDS) was then carefully added to maintain a level top edge while the acrylamide set. After one hour the overlay buffer was removed and the stacking gel poured on top of the separating gel (2.3.3). Immediately prior to loading the gel, haemolymph samples from *Drosophila* were boiled for 2 minutes, cooled on ice and mixed by pipetting. Proteins were separated by electrophoresis at 60V overnight in 1x gel running buffer (2.3.3). Following electrophoresis, proteins were visualised using the ISS Pro-Blue staining system (supplied by Integrated Separation Systems, Enprotech). The gel was fixed for 1 hour in 12 % (w/v) TCA, pre-treated for 1 hour in Pretreatment solution and stained overnight in Pro-Blue Staining solution with gentle agitation. If necessary, background blue staining was cleared in 25 % (v/v) methanol before the gel was photographed.

#### 2.4.3.11 *lacZ* Assays

Following treatment with 20-hydroxyecdysone, ten flies from individually transformed lines were dissected in Ringer's Solution and the fat bodies or thoraces thoroughly homogenised in an Eppendorf with 500µl of ice cold Homogenisation Buffer (2.3.5) using a hand held motorised homogeniser. Samples were then centrifuged in a microfuge at 12Krpm for 5 minutes and 450µl of the supernatant retained. 400µl of the supernatant was pre-incubated at 37°C for 5 minutes before starting the enzyme reaction by adding 600µl of o-Nitrophenyl-β-galactopyranoside (ONPG, 2.3.5). The samples were incubated at 37°C and the colour reaction, representing β-galactosidase activity, was measured at OD<sub>420nm</sub> on a Pye Unicam PU 8800 UV/VIS Spectrophotometer, at 10 minute time intervals for 115 minutes. The protein concentration of each sample was determined using the Bradford method employed in the Bio-Rad Protein Assay (2.4.4.2). 750µl of sterile water was added to the remaining 50µl of homogenate, and to this was added 200µl of Dye Reagent Concentrate. Samples were mixed by gentle inversion and the OD<sub>595nm</sub> measured after 10 minutes.

## **2.4.4 GEL RETARDATION ASSAYS**

### **2.4.4.1 Preparation of Ovarian Nuclear Extracts**

Ovarian Nuclear Extracts were essentially prepared as described by Frank et al., 1982. The solutions required are detailed in section 2.3.7. All stages of the procedure were carried out working either on ice or at 4°C. Approximately 400-600µl of ovaries were dissected in Ringer's Solution from healthy, egg-laying, 2-4 day old Oregon R females as these contained a high proportion of stage 8-10 oocytes. Ovaries were stored in 1 ml of MTBS Buffer, and could be stored overnight on ice in this buffer if two days were required to dissect sufficient ovaries. The ovaries were rinsed twice with MTBS to remove food particles and the egg chambers were gently separated, taking care not to break them, by careful pipetting with a 1ml Gilson pipette and a cut off blue tip. Egg chambers were collected by centrifugation at 2.2Krpm for 7 minutes at 4°C. The supernatant was removed and the egg chambers washed twice with 1ml of MTBS. Following the last centrifugation and removal of the supernatant, it was checked that the egg chambers were intact using a microscope. Washed egg chambers were then resuspended in 2 volumes of Solution I and thoroughly homogenised using a hand held motorised homogeniser. The nuclei, with cell debris, were collected by centrifugation at 4Krpm for 4 minutes at 4°C. A small volume of the pellet was checked under the microscope to ensure that the majority of egg chambers were broken, and the homogenisation step was repeated if necessary. The pellet was resuspended in an equal volume of Solution II. 5M NaCl was added to a final concentration of 0.35M and the sample was incubated on ice for 30 minutes. Incubation in 0.35M salt is done to strip most non-histone DNA-binding proteins from the DNA. The nuclear suspension was cleared of cell debris by centrifugation at 17Krpm overnight at 4°C. Glycerol was added to the supernatant to a final concentration of 20 % (v/v) and the extract was divided into 20µl aliquots for storage at -70°C. The protein concentration of the extract was determined using the Bio-Rad Bradford Assay (a yield of 2-4µg/µl was required for the gel retardation assays).

#### **2.4.4.2 Estimation of Protein Concentration of *Drosophila* Ovarian Nuclear Extracts**

Protein concentrations of 1-25µg/ml were estimated using the Bradford method (Bradford, 1976; Read and Northcote, 1981), available as a Protein Assay kit (Biorad). This involves the protein binding properties of Coomassie blue G-250. Binding causes a change in colour from red/brown to blue corresponding to an absorbance shift from 465nm to 595nm. 1µl and 10µl samples from the ovarian nuclear extract were added to 800µl of sterile water and to this was added 200µl of Dye Reagent Concentrate (contains Coomassie blue G-250, phosphoric acid, methanol and water). Samples were mixed by gentle inversion and the OD<sub>595nm</sub> measured between 5 minutes and one hour against a reagent blank. The concentration of protein in the sample was estimated from a standard graph of 3.125µg, 6.25µg, 12.5µg and 25µg BSA standards versus OD<sub>595nm</sub>. A standard graph was prepared each time the protein concentration of an extract was measured.

#### **2.4.4.3 Radioactive End-labelling of Probe**

200ng of digested, gel-purified DNA fragment (2.4.2.6b) was incubated in a total volume of 30µl with 6µl of Klenow 5x Buffer (0.25M Tris-HCl pH 7.2, 50mM MgSO<sub>4</sub>, 0.5mM dithiothreitol, 3µl of nucleotide solution containing 20mM dTTP, dGTP and dATP, 2µl (20µCi) α-<sup>32</sup>P[dCTP] and 1µl of *E.coli* DNA polymerase I Klenow fragment (5u/µl). The labelling reaction was allowed to proceed for 30 minutes at room temperature after which 5µl of 5mM dCTP was added and the reaction allowed to continue for a further 30 minutes at room temperature. The reaction volume was made up to 50µl with the addition of 15µl of TE buffer, and the radiolabelled fragments separated from unincorporated nucleotides by passing through a 1ml Sephadex G-50 NICK<sup>TM</sup> Column (detailed in section 2.4.3.4). The eluted probe was collected in 400µl of TE buffer, ethanol precipitated (2.4.2.2) at -20°C overnight and resuspended in 20µl TE buffer to give a probe concentration of approximately 10ng/µl (with radioactivity measuring at 10-30cps/µl).

#### **2.4.4.4 DNA-binding Reactions**

Ovarian nuclear extract (approximately 3µg protein) was incubated with 10-20ng of <sup>32</sup>P end-labelled DNA fragment (or probe), salmon testes carrier DNA (100ng-1µg), poly[d(I-C)] (10ng-1µg), and varying amounts of specific or non-specific competitor DNA (see figures 4.2A and B, 4.3A and B, 4.4A and B), made up to 20µl with Incubation Buffer (2.3.7). The binding reactions were carried out on ice for 30 minutes. 5µl of 5x gel shift loading buffer was added to each reaction and immediately loaded onto nondenaturing polyacrylamide gels for electrophoresis.

#### **2.4.4.5 Nondenaturing Polyacrylamide Gel Electrophoresis of Protein-DNA Binding Reactions**

The entire volume of the binding reactions were electrophoresed on a 1mm, 4 % nondenaturing high ionic strength (Tris-glycine) polyacrylamide gel containing 2.5 % glycerol, in high ionic strength electrophoresis buffer (2.3.7). Gels were pre-run for 1-2 hours at 150V during which time the current dropped from 35mA to 25mA. Gels containing binding reactions with a 150 bp probe were run for 2 hours at 150V, and those with ~300 bp probe for 2 hours and 30 minutes at 150V. Following electrophoresis, gels were rinsed in water, transferred onto blotting paper and dried under vacuum at 80°C for 1-2 hours. Gels were finally autoradiographed at room temperature overnight, or longer if required.

#### **2.4.4.6 Autoradiography**

Radioactive signals from dried DNA sequencing gels, dried protein gels and nylon membrane filters were detected using Dupont Cronex-4 X-ray film and cassettes with intensifying screens at -70°C for <sup>32</sup>P, and room temperature for <sup>35</sup>S. Films were developed in an Agfa 1 automatic film processor.

## **CHAPTER 3**

### **Results**

#### **Identification of *Cis*-acting Sequences Regulating *yp3* Transcription in *Drosophila melanogaster***

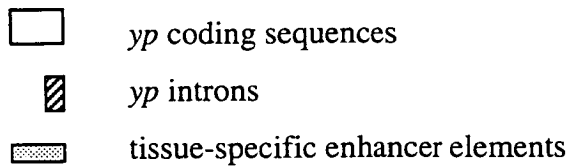
### 3.1 Introduction

In *Drosophila melanogaster*, the three *yolk protein (yp)* genes are transcribed in a sex-, tissue- and developmentally-specific manner, and therefore provide an ideal system in which to investigate the factors involved in their regulation. The yolk proteins (YPs) are encoded by three single copy genes on the X chromosome (Postlethwait and Jowett, 1980; Barnett et al., 1980), all of which have been cloned and sequenced (Hovemann et al., 1981; Hung and Wensink, 1981; 1983; Garabedian et al., 1987; Yan et al., 1987). The divergently transcribed *yp1* and *yp2* genes are located at 8F-9A and separated by a shared intergenic region of 1226 bp (Hung and Wensink, 1983), whereas *yp3* is isolated with separate regulatory sequences at 12BC (Garabedian et al., 1987) (Figure 3.1). Nevertheless, all three genes are coordinately regulated such that transcription occurs only in the fat body of adult females and in the ovarian follicle cells, that surround the developing oocyte, during stages 8-10 of oogenesis (Bownes and Hames, 1978; Brennan et al., 1982; Isaac and Bownes, 1982). The mechanisms involved in this regulation are therefore thought to involve sex- and tissue-specific factors that must be present at the correct developmental stages for gene activation.

Investigations into *yp* gene control have mainly focused on *yp1* and *yp2*, for which several *cis*-acting elements have been identified that are necessary for their correct pattern of expression (Figure 3.1). This is discussed in detail in section 1.5 of the Introduction. In summary, a 125 bp fat body element (FBE) located 196 bp upstream of the *yp1* transcription start site was identified that was sufficient to direct the sex-specific fat body expression of both genes (Garabedian et al., 1986). However, it was later shown that other sequences within the 1226 bp intergenic region were equally capable of conferring female-specific *yp* transcription in the fat body when the FBE was deleted (Abrahamsen et al., 1993). Ovarian expression of *yp1* and *yp2* was discovered to be regulated by two tissue-specific enhancers, ovarian enhancer 1 (OE1) and ovarian enhancer 2 (OE2). OE1 was a 301 bp sequence located 43 bp upstream of the *yp2* transcription start site required to direct the correct developmental profile of *yp* transcription. Expression of the *yps* in the correct

### Figure 3.1

The genomic arrangement of the *yp* genes with their location on the X chromosome



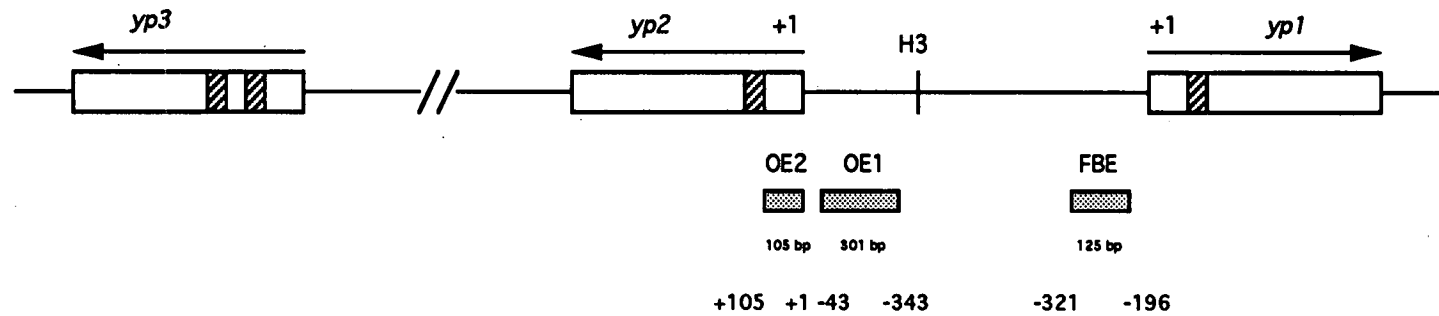
Transcripts are represented by arrows, with the arrowhead representing the direction of transcription. The locations and sizes of the fat body enhancer (FBE) and ovarian enhancers (OE1 and OE2) are also shown. Nucleotide positions associated with the FBE are given relative to the *yp1* transcription start site and the nucleotide positions associated with OE1 and OE2 are given relative to the *yp2* transcription start site



LOCATION ON  
X CHROMOSOME

12BC

8F - 9A



follicle cell types was found to result from the combined positive and negative effects of OE1 regions on *yp* expression in different subpopulations of follicle cells at stages 9-11 of oogenesis. However, OE1 was not sufficient on its own to maintain high levels of ovarian *yp1* transcripts and this led to the identification of OE2 which was located in the first 105 bp of the *yp2* coding region. OE2 acted across the *yp2* promoter to enhance levels of ovarian *yp1* transcripts. In the absence of OE1, OE2 was sufficient to limit *yp1* expression to the ovarian follicle cells but this was not restricted to the usual stages 8-10 of oogenesis. Both ovarian enhancers were therefore required to maintain correct levels of *yp1* and *yp2* transcription in the ovarian follicle cells during limited stages of oogenesis (Logan et al., 1989; Logan and Wensink, 1990).

From these data, it is apparent that an understanding of the control of *yp1* and *yp2* transcription is complicated by their shared *cis*-acting regulatory elements. Several regions within the intergenic region are able to direct the sex-specific fat body expression of *yp1* and *yp2* and two ovarian enhancers have been identified that are required for the correct ovarian pattern of *yp* expression, one of which is located in the first exon of the *yp2* coding sequence. Further insight into the mechanisms governing *yp* transcription will require the isolation and identification of *trans*-acting regulatory proteins, and an analysis of their mode of action, which firstly requires the precise identification of the *cis*-acting elements with which they interact.

For this reason, most of the transcriptional regulatory proteins implicated in the control of *yp* sex-limited, tissue-specific expression have been identified by *in vitro* footprinting assays with the *yp1/yp2* FBE. For example, the male- and female-specific doublesex proteins (DSX<sup>M</sup> and DSX<sup>F</sup>) encoded by the differentially spliced *dsx* gene, the last gene in the *Drosophila* sex determination pathway, are known to be involved in the regulation of *yp* expression since *dsx* must be expressed normally for *yp* expression to occur in the female fat body (Shirras and Bownes, 1987; Slee and Bownes, 1990; Steinmann-Zwicky et al., 1990). Footprinting assays have been used to demonstrate that both DSX<sup>M</sup> and DSX<sup>F</sup> bind to the same sequences in the *yp1/yp2* FBE. Four DSX binding sites were reported to exist (Burtis et al., 1991), three of which were mutated and the effects monitored by

*in vivo* analysis and DNA binding *in vitro*. Results from these experiments suggested that DSX<sup>M</sup> repressed and DSX<sup>F</sup> activated *yp* expression from the two strongest binding sites. Also, if both proteins were present in binding studies, they competed for the same binding sites (Coschigano and Wensink, 1993).

*Trans*-acting factors thought to be involved in regulating the tissue-specificity of *yp* expression in the fat body are the transcriptional activator box B-binding factor-2 (BBF-2), the CCAAT/enhancer binding protein (C/EBP) and the adult enhancer factor-1 (AEF-1). These proteins were initially found to interact with the *Drosophila melanogaster Adh* adult enhancer to regulate *Adh* expression in the adult fat body. More recently, footprinting studies have shown that they also bind to sites within the *yp1/yp2* FBE (Abel et al., 1992; Falb and Maniatis, 1992a, 1992b). It is thought that AEF-1 and C/EBP compete for their binding sites, which overlap in the FBE, since AEF-1 can suppress the activating potential of C/EBP by preventing it from binding to its target sequence (Falb and Maniatis, 1992b).

These factors have all been shown to bind to sequences within the *yp1/yp2* FBE (represented in Figure 3.9A) yet the FBE can be deleted from the *yp1/yp2* intergenic region without losing sex-limited, tissue-specific *yp* expression (Abrahamsen et al., 1993). This suggests that there is some redundancy of the *cis*-acting sequences required for the maintenance of the *yp* expression pattern and that other sequences in the intergenic region are equally capable of regulating *yp* transcription. Although computer searches with the target sequences for DSX, BBF-2, AEF-1 and C/EBP have located several alternative potential binding sites for these factors in the intergenic region (Bownes et al., 1993) the significance of these need to be investigated by footprinting assays, mutational analysis, and the effects monitored *in vivo*.

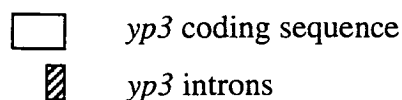
There have been no reports on the isolation of regulatory proteins that govern tissue-specific ovarian transcription. A possible candidate for this was yolk protein factor 1 (YPF1), a heterodimeric DNA-binding protein reported to bind with high affinity to a 31 bp fragment located 148 bp downstream of the *yp1* transcription start site. This protein was expressed in *Drosophila* throughout development, and high levels of transcripts were found in the oocyte and early embryos (Mitsis and

Wensink, 1989a, 1989b). However, YPF1 has recently been reported to be a *Drosophila* homologue of Ku, a DNA-dependent protein kinase from humans. Ku and YPF1 are therefore thought to have multiple functions and could be involved in transcription initiation, recombination or repair (Jacoby and Wensink, 1994). Thus, YPF1 does not appear to be involved in the control of tissue-specific ovarian *yp* transcription and the factors regulating this are yet to be isolated. The isolation of ovarian *trans*-acting factors involved in regulating *yp1* and *yp2* ovarian transcription has possibly not been pursued because no single *cis*-acting element is responsible for their expression in this tissue.

Although *yp3* is located some distance away from *yp1* and *yp2* on the X chromosome, all three *yp* genes have the same highly specific and coordinated pattern of expression. It is therefore likely that the *cis*- and *trans*-acting factors involved in regulating their transcription will be similar, if not the same. However, it was anticipated that the *cis*-acting factors involved in the control of *yp3* transcription would be more straightforward to identify as they would not be shared by neighbouring genes. Preliminary analysis with *yp3* had shown that a 4.3 kb genomic fragment consisting of the *yp3* coding sequence, 1822 bp of 5' flanking DNA and 825 bp of downstream sequence was sufficient to direct the correct pattern of *yp* expression in *yp3*<sup>-</sup> flies (Figure 3.2A). A similar construct, that only differed in that it contained 704 bp of 5' flanking DNA (Figure 3.2B) was also shown to be sufficient to direct the correct pattern of *yp* expression. *yp3* expression in the fat body was detected by analysing haemolymph samples from transformed flies on Coomassie Blue stained SDS-PAGE gels. YP3 synthesis in the ovary was verified by culturing the ovaries from transformed females and analysing the proteins in the culture medium by Western blotting. These experiments demonstrated that *yp3* had been isolated with its own regulatory sequences since both constructs were sufficient to direct sex-limited, tissue-specific YP synthesis in the ovarian and fat body tissue of adult females (Liddell and Bownes, 1991). The initial aim of this project was to further separate and precisely locate the *cis*-acting regions required for the sex- and tissue-specific pattern of *yp3* expression.

### Figure 3.2

Diagrammatic representation of the *yp3* constructs tested for transcription enhancer elements and their ability to direct reporter gene expression in the ovary and/or fat body in transformed flies. At the top is a schematic diagram of the *yp3* gene with the restriction enzyme sites used to create the different reporter constructs



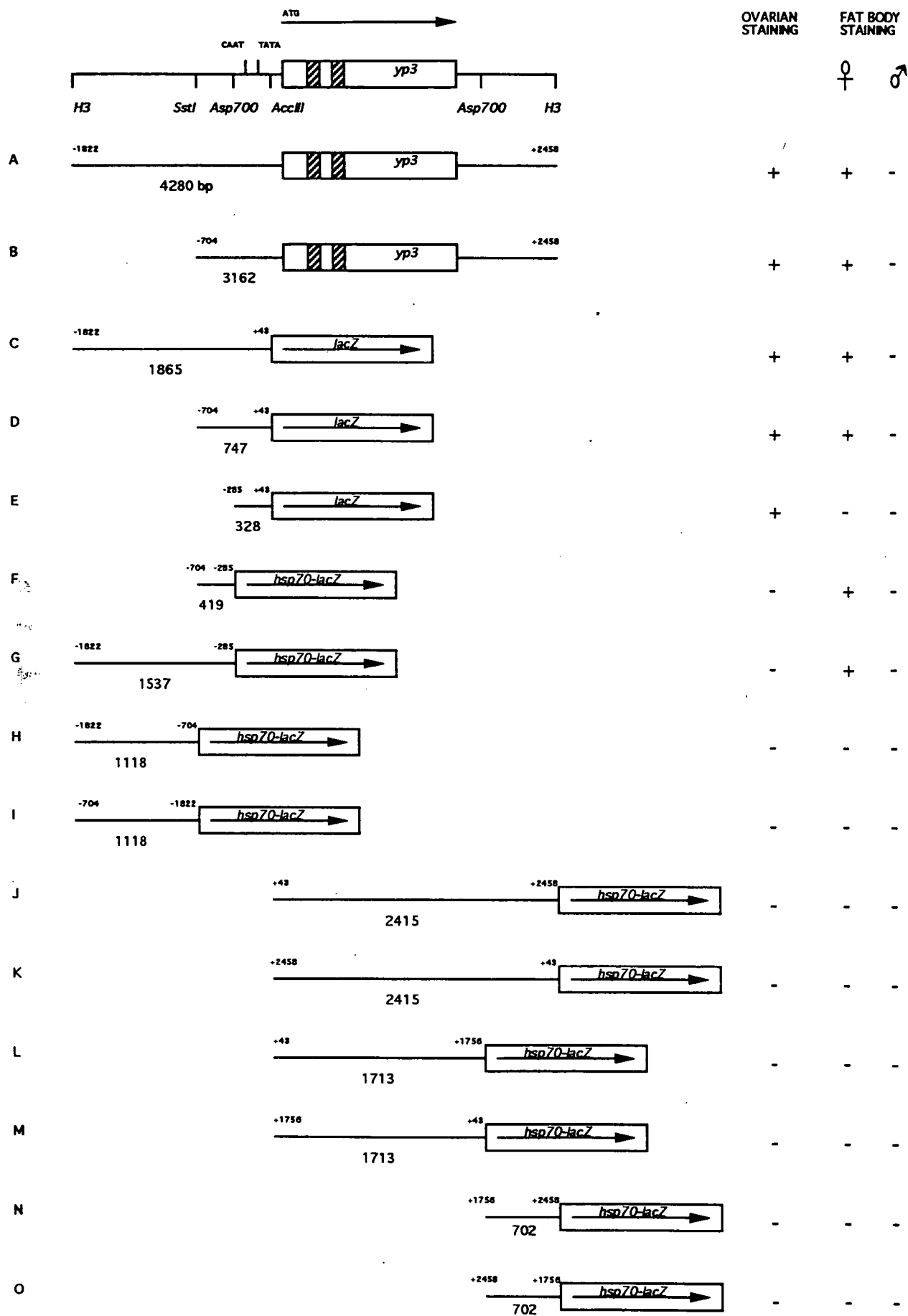
Transcripts are represented by arrows and the arrowhead represents the direction of transcription. The sizes (in bp) of the *yp3* fragments used in reporter constructs are indicated below the constructs and the numbers in small print, above the constructs, represent the nucleotide positions in relation to the *yp3* transcription start site and the orientation of the fragment with respect to the reporter gene.

The ability of all constructs to direct reporter gene expression was assessed by  $\beta$ -galactosidase activity, with the exception of constructs A and B, where expression was confirmed by the presence of YP3 protein on SDS-PAGE gels and Western blots (Liddell and Bownes, 1991). Constructs C-E, that contained *yp3* TATA and CAAT-box promoter sequences, were cloned upstream of the *lacZ* reporter gene contained in the pCaSpeR-AUG- $\beta$ gal transformation vector (Figure 3.3A) and constructs F-O, lacking the *yp3* promoter region, were cloned into this vector modified to contain a *hsp70-lacZ* gene fusion (Figure 3.3B)

+ indicates  $\beta$ -galactosidase activity

- indicates no reporter gene expression

*Handwritten signature*



### 3.2 Analysis of the *yp3* Gene to Identify *Cis*-acting Elements Required for Sex-limited, Tissue-Specific Expression

The upstream, downstream and coding regions of *yp3* were separated in order to further localise any tissue-specific enhancer elements present. It was of interest to determine whether *yp3* was regulated by a number of elements that could act independently, as occurred with the sex-specific fat body expression of *yp1* and *yp2* (Logan et al., 1989; Abrahamsen et al., 1993) or whether separate elements were interacting to coordinate expression such as OE1 and OE2 in the regulation of *yp1* and *yp2* ovarian expression (Logan et al., 1989).

The *yp3* gene was separated into various 5', coding and 3' components using the suitable restriction sites shown in Figure 3.2. Fragments containing the *yp3* promoter (Figure 3.2, constructs C, D and E) were fused upstream of the *E.coli lac Z* ( $\beta$ -galactosidase) reporter gene contained within the P-element transformation vector pCaSpeR-AUG- $\beta$ gal (Thummel et al., 1988). The remaining constructs (F-O) were tested in a modified version of this vector containing sufficient *Drosophila hsp70* promoter sequences for gene activation in the presence of enhancers.

#### 3.2.1 The pCaSpeR-AUG- $\beta$ gal Transformation Vector

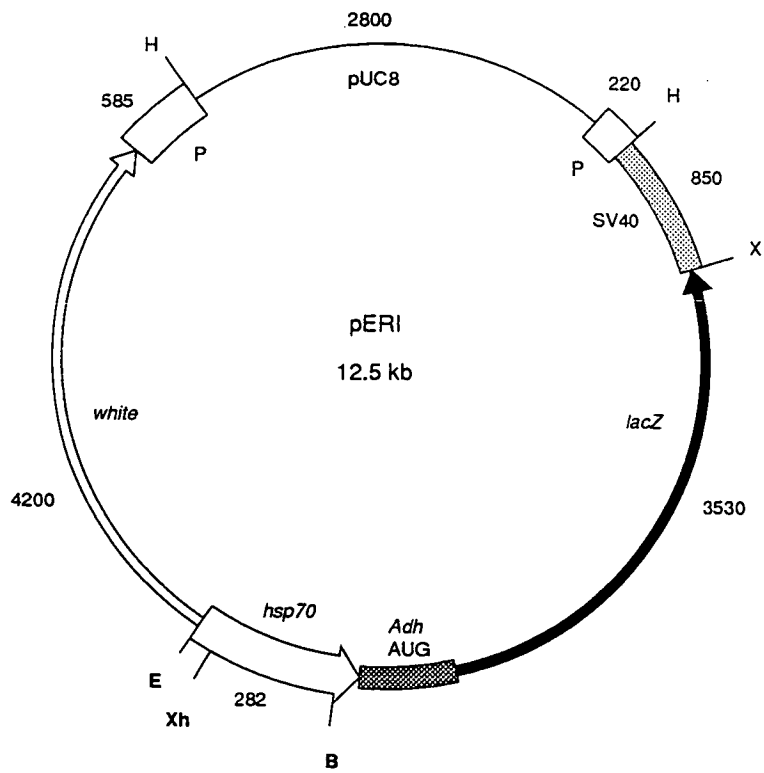
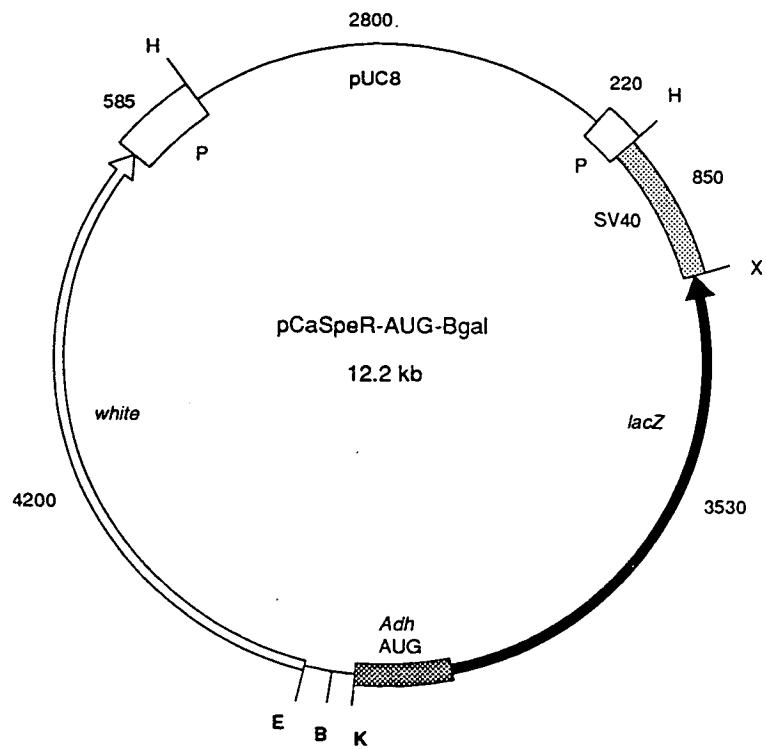
The pCaSpeR-AUG- $\beta$ gal transformation vector (Figure 3.3A) contains the *Drosophila white*<sup>+</sup> gene as a genetic marker, such that transformants can be identified by screening for restoration of eye pigmentation in *white*<sup>-</sup> flies (Lüning, 1981). Three unique restriction sites, *EcoRI*, *BamHI* and *KpnI* are available upstream of the *lacZ* reporter gene for the insertion and testing of foreign promoters. The *lacZ* gene also has an AUG start codon from the *Drosophila Adh* gene fused upstream of the open reading frame such that promoters to be tested do not have to be fused in frame with the reporter gene. A polyadenylation signal to allow appropriate post-transcriptional RNA processing of the *lacZ* reporter gene is provided by SV40 DNA. The *white* gene, multiple cloning site, *Adh* AUG sequence,

### Figure 3.3

**A** Structure of the pCaSpeR-AUG- $\beta$ gal transformation vector. Unshaded boxes at the *Hind*III sites (H) indicate *P*-element ends. The *white* gene is transcribed in the opposite direction to the *lacZ* reporter gene (arrows show the direction of transcription) and unique cloning sites are *Eco*RI (E), *Bam*HI (B) and *Kpn*I (K), shown in bold. The *Adh*-AUG sequence ensures that all reporter constructs are in-frame fusions to the *lacZ* gene. A trailer region derived from SV40 on a *Hind*III (H) - *Xba*I (X) fragment provides the *lacZ* polyadenylation signal. The size of the vector is 12.2 kb and the sizes (in bp) of the vector components are shown

**B** Structure of the pERI transformation vector. The pCaSpeR-AUG- $\beta$ gal vector was modified to contain a 282 bp *hsp70* promoter fragment. Unique cloning sites are now *Eco*RI (E) and *Xho*I (Xh) upstream of the *hsp70* sequences, and *Bam*HI downstream of the *hsp70* promoter. The size of the vector is 12.5 kb





*lacZ* reporter gene and SV40 DNA are flanked by P-element sequences for insertion into the *Drosophila* genome following injection of the vector into *Drosophila* embryos.

### 3.2.2 The pERI Transformation Vector

In order to test *yp3* constructs lacking promoter sequences (Figure 3.2, constructs F-O) for the presence of enhancer elements, the pCaSpeR-AUG- $\beta$ gal vector was modified by cloning a 282 bp *XhoI-PstI* fragment from the *Drosophila hsp70* promoter region upstream of the *Adh*-AUG region (Figure 3.3B). This *hsp70* promoter fragment contained sequences reported to be necessary for transcriptional activation of the *hsp70* gene (Figure 3.4). These included a TATAATA eukaryotic promoter sequence located 23 bp upstream from the proposed 5' end of the mRNA for the 70 kDa heat shock protein, identical to the consensus TATA-sequence found 20-25 bp upstream of the transcription initiation site in a wide variety of genes. This was flanked by G-C rich regions, which are thought to be involved in linking distant control elements (Briggs et al., 1986), the first 5 bp of which form an inverted repeat of the sequence CGGAG. A sequence similar to the mRNA cap site (CAATTCA) occurred twice immediately following the proposed 5' end of the mRNA. A dodecanucleotide sequence of TGTTCGCGAAAA was identified 12 bp upstream from the start of the A-T rich sequence that had dyad symmetry with one mismatch. This was thought to be involved in controlling the expression of the *hsp70* gene as symmetrical sequences are often associated with protein-DNA interactions involved in regulating gene transcription (Ingolia et al., 1980). It was later shown that this sequence was part of a neighbouring sequence, CTCGAATGTTCGCGA, necessary for heat inducible gene activity. This was demonstrated by assessing the effects of 5' deletions on the levels of *hsp70* transcripts in COS cells (SV40 transformed monkey cells that support high levels of replication) and *Xenopus* oocytes (Pelham, 1982). A consensus 14 bp symmetrical sequence CTnGAAnnTTCnAG, present in nearly all heat shock genes, was shown to be necessary for heat shock induction of transcription by placing synthetic oligonucleotides upstream of the TATA box of the

### Figure 3.4

**A** The *XhoI-PstI hsp70* promoter fragment cloned upstream of the *Adh*-AUG region in pCaSpeR-AUG- $\beta$ gal, to create the pERI transformation vector

**B** The sequence between nucleotides -89 and +23 containing promoter elements known to be required for transcriptional activation of the *hsp70* gene. This is discussed in detail in the text. The TATA sequence is highlighted in bold. The mRNA cap site is underlined and immediately follows the transcription initiation site (+1), outlined. The G-C rich inverted repeat regions, flanking the TATA sequence, are shown in lower case letters and the sequences located at nucleotide positions -46 (required for heat shock induction of transcription) and -73 (required for normal levels of transcription following heat shock) are printed underneath the sequence.

All nucleotide positions are given relative to the transcription start site (+1) of the *hsp70* gene

A



B

-89                      -73                      -61                      -46                                      -23                                      +1                                      +23

TACTGCTCTCGTTGGTTCGAGAGAGCGCGCCTCGAATGTTTCGCGAAAAGAGCGCcgagTATAAATAgaggcGCTTCGTCTACGGAGCGACAATTCAATTCAAACAAGCAAA

GTTGGTTCG                      CTnGAA nnTTnAG

herpes virus thymidine kinase (*tk*) gene. This sequence conferred heat shock induction of *tk* gene expression in COS cells and *Xenopus* oocytes (Pelham and Bienz, 1982). This information was confirmed in *Drosophila*, with the establishment of P-element mediated germ-line transformation techniques, by studies with 5' deletions of the *hsp70* promoter region. Constructs were fused to a truncated *Drosophila Adh* gene lacking its own promoter and transformed back into flies to study the effects. It was discovered that 97 bp of 5' flanking DNA was sufficient to induce transcription upon heat shock, to a similar level achieved with the wild-type *hsp70* gene. However levels of reporter gene induction were comparably low in flies transformed with constructs containing 68 bp of upstream sequence, even though this contained the 14 bp symmetrical sequence shown to be sufficient to induce transcription upon heat shock in COS cells and *Xenopus* oocytes. A second promoter element was therefore required in *Drosophila* for normal levels of transcription following gene activation by heat shock, that was located between nucleotide positions -68 and -97 upstream of the *hsp70* transcription start site (Dudler and Travers, 1984). These observations were confirmed by an independent analysis where similar *hsp70* promoter constructs were fused upstream of the *lacZ* reporter gene (Amin et al., 1985). The sequence GTTGGTTCG located 73 bp upstream of the *hsp70* transcription initiation site is probably the sequence required for normal levels of heat shock induced transcription. This was identified by Ingolia et al. (1980) as a possible sequence involved in the transcriptional regulation of the *hsp70* gene as it had partial homology to a consensus recognition sequence found 70-80 bp upstream of several eukaryotic genes, thought to be involved in modulating the rate of gene transcription.

The *XhoI-PstI* fragment was firstly cloned into the polylinker of pBluescript SK (Stratagene, Inc.) and then into pUC19 on a *KpnI-XbaI* fragment. The *hsp70* construct could then be incorporated into the pCaSpeR-AUG- $\beta$ gal vector in the correct orientation, using the *EcoRI* and *BamHI* restriction sites.

The resulting transformation vector, pERI, contained all the *hsp70* promoter sequences shown to be necessary for correct transcriptional induction of gene expression following heat shock, and should therefore allow fragments lacking

the *yp3* promoter (Figure 3.2, constructs F-O) to be tested for their effects on reporter gene activity. Unique restriction sites in pERI for cloning of the constructs to be tested were *EcoRI* and *XhoI* positioned 5' of the *hsp70* promoter, and *BamHI* positioned 3' of the promoter (Figure 3.3B). The pERI transformation vector was tested to ensure that there were no background levels of *lacZ* expression and that the *hsp70* promoter sequences were operational. Transformants were heat shocked by incubation at 37°C for 45 minutes and allowed to recover for 1 hour at room temperature prior to dissection and histochemical staining. No enzyme activity was detected unless transformants were subjected to heat-shock, in which case staining was observed in all tissues of both males and females, demonstrating that the *hsp70* promoter sequences present were functional and sufficient to activate gene expression. Photographs demonstrating this are shown in Figure 3.5.

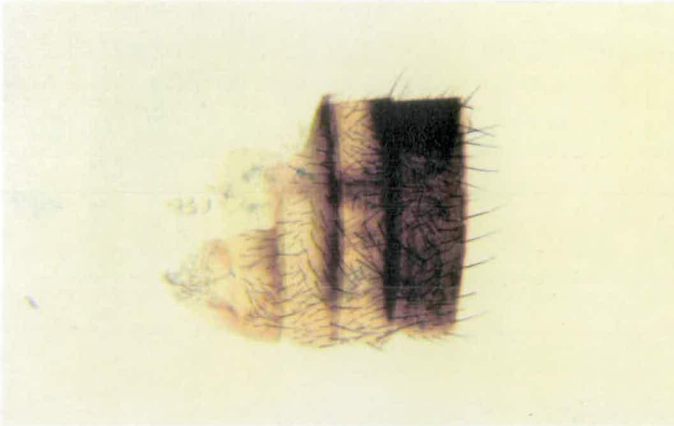
### 3.2.3 Construction and Analysis of the *yp3* Reporter Constructs

The *yp3* regions tested for the presence of enhancer elements are illustrated in Figure 3.2. Construct C was a 1865 bp fragment (-1822 to +43) containing the entire 5' flanking region that was isolated with the 4.3 kb *yp3* genomic fragment. It also contained 43 bp of transcribed but non-translated sequence. Construct D consisted of a smaller upstream region of 747 bp (-704 to +43). Therefore constructs C and D were similar to constructs A and B respectively with regard to the amount of upstream sequences present. The only difference was that the expression directed by constructs C and D would be detected by  $\beta$ -galactosidase activity rather than, as was the case for constructs A and B, the synthesis of YP3 in a YP3<sup>-</sup> background. This would also give a preliminary indication of whether enhancer elements associated with *yp3* were present in the upstream sequences, or in the coding and/or downstream regions. Construct D was divided into two further smaller constructs, E and F of 328 bp (-285 to +43) and 419 bp (-704 to -285) respectively. Construct E contained sequences immediately upstream of the *yp3* coding sequence that included both the TATA and CAAT-box promoter elements. Construct G was a 1537 bp fragment spanning nucleotides -1822 to -285. It was

### Figure 3.5

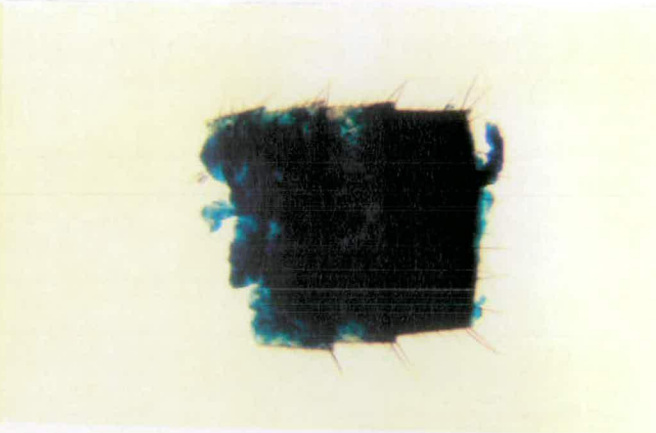
Photographs representing  $\beta$ -galactosidase activity in the abdominal fat body of males and females transformed with the pERI vector prior to and following heat shock. No staining was observed in transformed males or females without heat shock (**A** and **C**) indicating very low background levels of reporter gene expression. Following heat shock, staining was observed in both males and females in all tissues, although only the fat body, epidermis and muscle staining in the abdomen is shown here (**B** and **D**). This demonstrated that the *hsp70* promoter sequences were sufficient to activate reporter gene expression

A



MALE ABDOMEN - HEAT SHOCK

B



MALE ABDOMEN + HEAT SHOCK

C

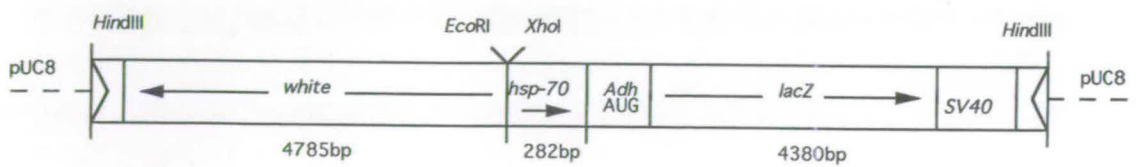


FEMALE ABDOMEN - HEAT SHOCK

D



FEMALE ABDOMEN + HEAT SHOCK





therefore composed of the *yp3* upstream region, only lacking the region containing the *yp3* promoter sequences. Constructs H and I were the final upstream sequences to be tested, in a forward and reverse orientation respectively (with respect to the direction of transcription), and spanned nucleotides -1822 to -704. The *yp3* gene was also divided into its coding and 3' regions, which were tested together and separately to investigate whether this gene could be involved in regulating its own transcription. Constructs J and K, spanning nucleotides +43 to +2458, consisted of the coding and downstream sequences together, tested in a forward and reverse orientation respectively. Constructs L and M were 1713 bp in length, spanning nucleotides +43 to +1756, and consisted of the *yp3* coding sequence with 150 bp of 3' flanking sequence tested in a forward and reverse orientation. Finally, 702 bp of 3' flanking DNA was tested in both orientations, as depicted by constructs N and O in Figure 3.2. All nucleotide positions are given in relation to the *yp3* transcription start site.

Constructs C, D and E containing *yp3* promoter sequences were cloned into the pCaSpeR-AUG- $\beta$ gal transformation vector. Construct C was isolated by digestion with *AccIII* and recessed 3' ends were filled in using Klenow, before *HindIII* digestion. This fragment was cloned into pBluescript SK cut with *XhoI*, Klenow treated to create blunt ends, and *HindIII* digested. It was then directionally cloned into pCaSpeR-AUG- $\beta$ gal on an *EcoRI-KpnI* fragment. A similar approach was used to clone construct D. *AccIII* digestion followed by Klenow end-filling and *SstI* restriction allowed this fragment to be cloned firstly into pUC19 digested with *SstI* and *SmaI*. The construct was then cloned into pCaSpeR-AUG- $\beta$ gal on an *EcoRI-BamHI* fragment. Construct E was isolated by digestion with *Asp700* and *AccIII*. The *AccIII* site was blunt-ended and *EcoRI* linkers ligated to this fragment for cloning into the pCaSpeR-AUG- $\beta$ gal transformation vector. Constructs F-O, lacking *yp3* promoter sequences, were isolated using the appropriate restriction sites and cloned into the *EcoRI* site of the pERI transformation vector by the creation of blunt ends followed by the addition of *EcoRI* linkers. All construct orientations were verified by dideoxynucleotide sequencing (2.4.2.9).

All constructs were introduced into the *Drosophila* germ line using standard transformation techniques (Rubin and Spradling, 1982; Spradling and

### Figure 3.6

Example of Southern blot of *Eco*RI digested genomic DNA extracted from *white*<sup>-</sup> lines independently transformed with constructs D (Di-Div), J (Ji-Jiii) and E (Ei-Eiv) illustrated in Figure 3.2. The probe used was a 850 bp *Xba*I-*Pst*I fragment from the transformation vector containing only SV40 sequences. The numbers on the right hand side indicate the size (in kilobase pairs) of the 1 kb ladder DNA marker (Bethesda Research Laboratories)

Di Dii Diii Div Ji Jii Jiii Ei Eii Eiii Eiv



- 12.2

- 9.1

- 7.1

- 6.1

- 5.1

Rubin, 1982) (2.4.3.1). Southern blot analysis was used to select independently transformed lines for further analysis that contained single copy inserts (2.4.3.3), an example of which is given in Figure 3.6. At least three independently transformed lines were assayed for each construct tested and reporter gene expression was detected by blue colouration derived from  $\beta$ -galactosidase activity, following histochemical staining of dissected ovaries and fat bodies (2.4.3.6).

### 3.3 Results and Conclusions

Of the thirteen constructs tested only two were able to direct reporter gene expression in both the female fat body and the correct stages of ovarian follicle cells. They were construct C, the 1865 bp fragment (-1822 to +43) containing the entire *yp3* upstream sequence, and construct D, the 747 bp fragment (-704 to +43) containing 704 bp of *yp3* upstream sequence. Photographs representing the staining observed in the fat body and ovaries of females transformed with construct D are given in Figure 3.7 C and D. Similar staining was observed in females transformed with construct C. These constructs were also sufficient to confer the sex-specificity of *yp* expression as no  $\beta$ -galactosidase activity was observed in the fat body tissue of males from the same transformed lines (Figure 5.4A). In the fat bodies of staged female transformants, *lacZ* expression was low following eclosion and increased gradually to reach a maximum after two to three days, reflecting the normal pattern of *yp* expression in this tissue (Barnett et al., 1980; D.Mauchline and M.Bownes, personal communication). This demonstrated that constructs C and D were able to direct a normal developmental profile of *yp* expression. In the ovaries,  $\beta$ -galactosidase activity was restricted to the ovarian follicle cells during stages 8 to 10 of oogenesis (Figure 3.7 D). The 5' region (-1822 to -704) absent from the shorter 747 bp fragment was independently tested for the presence of other enhancer elements, but no  $\beta$ -galactosidase activity was apparent in the fat body or ovaries of lines transformed with this construct (Figure 3.2, constructs H and I). Unfortunately, only one transformed line was obtained with this construct in the forward orientation (C.Simpson, personal communication). No *lacZ* expression was detected in any

### Figure 3.7

Photographs illustrating the  $\beta$ -galactosidase activity in female *Drosophila* transformed with *yp3-lacZ* or *yp3-hsp70-lacZ* fusion constructs, shown below the photographs. X-Gal stained abdomen (A) and dissected ovary (B) of flies transformed with the pERI vector (Figure 3.3B), showing an absence of staining or reporter gene activity. The oocyte (o), nurse cells (n) and follicle cells (f) are labelled. An example is shown of the fat body (C) and ovarian staining (D) for which similar patterns are seen in flies transformed with constructs containing either 1865 bp or 747 bp of *yp3* upstream sequence (Figure 3.2C and D). The fat body cells (fb) located under the abdominal cuticle (c) are stained, as are the ovarian follicle cells during stages 8-10 of oogenesis. Numbers indicate the developmental stages of egg chambers (King, 1970)

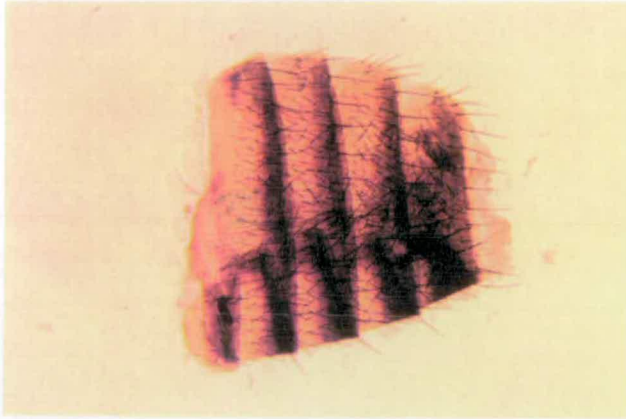
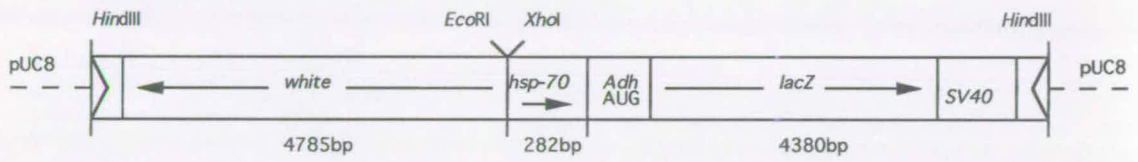
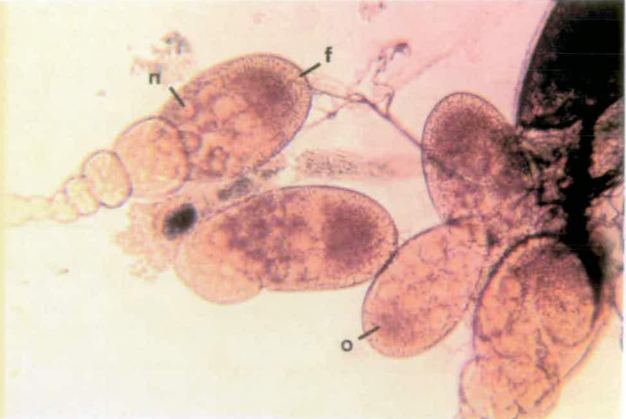
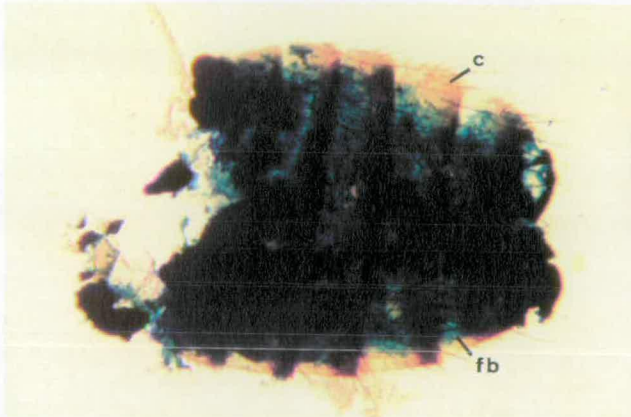
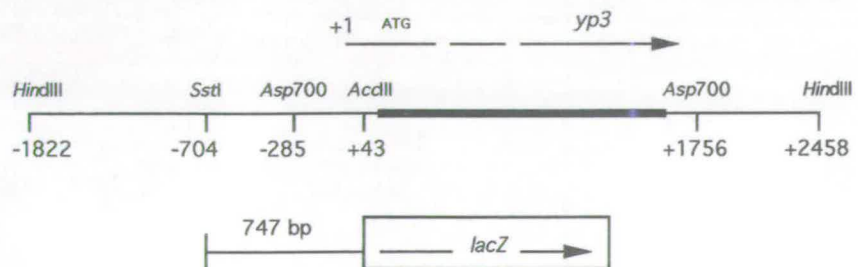
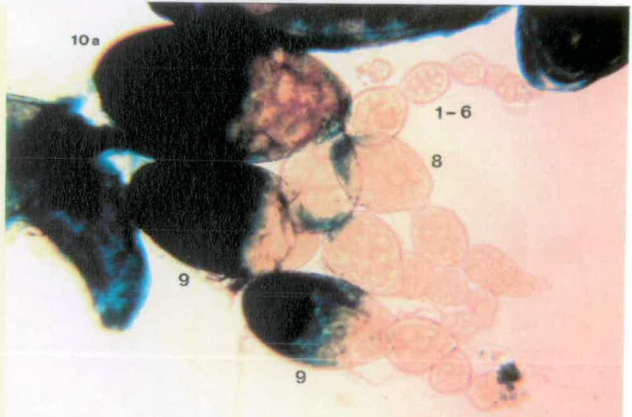
**A****FEMALE ABDOMEN****B****DISSECTED OVARY****C****FEMALE ABDOMEN****D****DISSECTED OVARY**

Figure 3.7 (cont.)

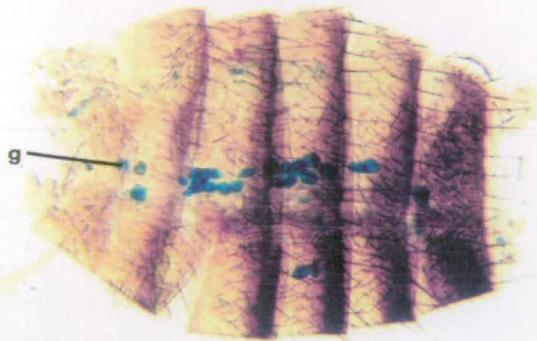
E and F      Photographs represent the staining patterns observed in the fat body and ovaries of flies transformed with 328 bp of *yp3* upstream sequence spanning nucleotides -285 to +43 (Figure 3.2E). The construct is shown below the photographs. Ovarian expression is maintained in the follicle cells of stage 8-10 oocytes. Numbers indicate the developmental stages of the egg chambers. There is no *lacZ* transcription in the fat body.

G and H      Photographs representing the staining patterns observed in the fat body and ovaries of flies transformed with the 419 bp *yp3* upstream fragment spanning nucleotides -704 to -285, illustrated below the photographs. Reporter gene expression and therefore  $\beta$ -galactosidase activity is seen in the fat body, but transcription in the ovaries is lost. A similar expression pattern is seen in flies transformed with the construct containing 1537 bp of *yp3* upstream sequence spanning nucleotides -1822 to -285 (Figure 3.2G)

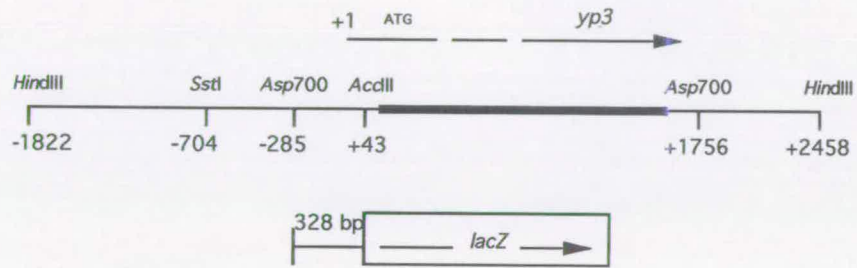
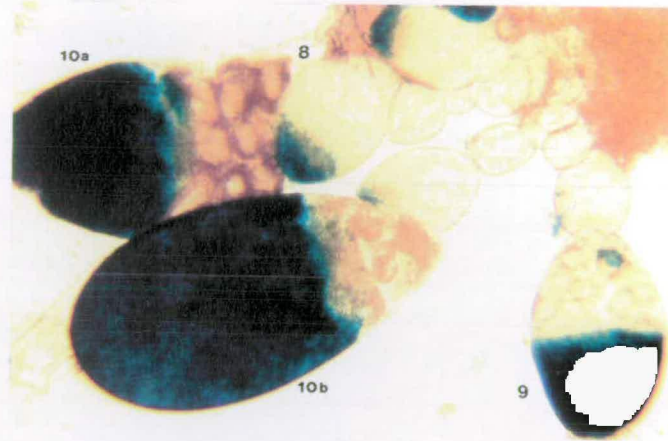


**E**

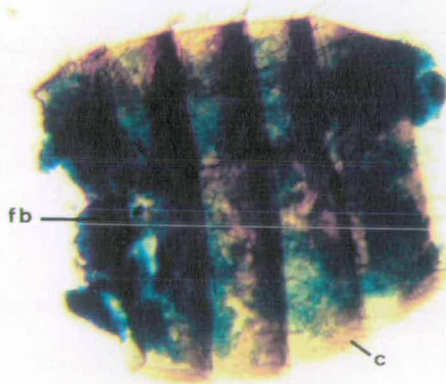
FEMALE ABDOMEN

**F**

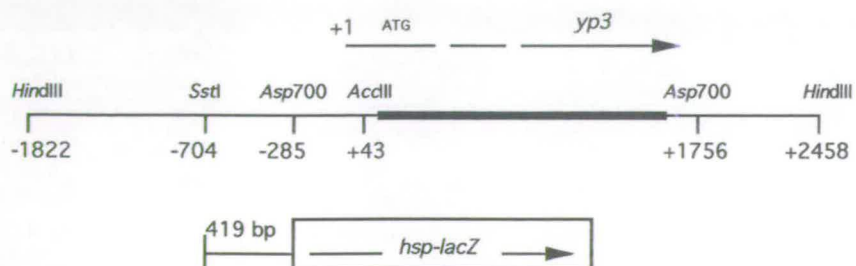
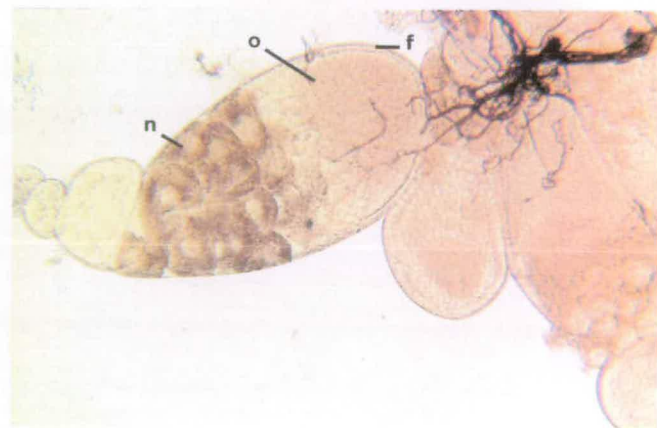
DISSECTED OVARY

**G**

FEMALE ABDOMEN

**H**

DISSECTED OVARY





tissues of males or females, although more lines would need to be assessed to eliminate the possibility that this resulted from a chromosomal position effect. Nevertheless, no other constructs were analysed that lacked expression in one particular line if they contained sequences responsible for directing tissue-specific gene transcription. Furthermore, although this result is mostly based upon this fragment when cloned in a reverse orientation upstream of the reporter gene, any enhancers present would, by definition, be expected to remain functional regardless of orientation (Serfling et al., 1985). The reporter constructs containing the *yp3* coding and downstream non-coding regions (Figure 3.2, constructs J-O) were also investigated for their ability to influence reporter gene expression. However, no staining was observed that resembled the *yp* expression pattern, suggesting that they were not involved in directing tissue-specific *yp3* transcription (data not shown). From this analysis it was concluded that the 747 bp region immediately upstream of the *yp3* coding sequence was the only region to contain *cis*-acting elements autonomously involved in controlling the highly regulated pattern of *yp3* transcription.

### **3.3.1 Isolation of *yp3* Upstream Sequences that Govern Tissue-Specific Ovarian and Sex-Limited Fat Body Expression of *yp3***

Two smaller constructs divided the 747 bp region into fragments of 328 bp (Figure 3.2, construct E) and 419 bp (Figure 3.2, construct F). The 328 bp fragment (-285 to +43) containing the *yp3* TATA and CAAT-box promoter elements, was sufficient to direct the correct pattern of ovarian *yp* expression, restricting reporter gene expression to the follicle cells of stage 8-10 oocytes. (Figure 3.7F). No staining was detected in the fat body or any other tissues of flies transformed with this construct (Figure 3.7E). The 419 bp fragment (-704 to -285) contained sequences able to direct sex-limited gene expression in female fat bodies (Figure 3.7G) since no fat body reporter gene expression was seen in males transformed with this construct. No staining was detected in the ovaries of females transformed with this construct (Figure 3.7H). This 419 bp fragment therefore contained sufficient

sequences to direct sex-specific fat body reporter gene expression from a heterologous promoter in adult females and was termed the *yp3*FBE. A similar result was seen in lines transformed with construct G that was 1537 bp and spanned nucleotides -1822 to -285. This contained the entire *yp3* upstream region apart from the 328 bp promoter fragment. This observation further demonstrated that ovarian expression was lost when the *yp3* promoter region was deleted. We therefore concluded that only the 328 bp fragment contained elements necessary for *yp* transcription in the ovarian follicle cells, and this region was termed ovarian enhancer for *yp3* or OE3.

During this analysis, it was noticed that constructs containing *hsp70* promoter fusions were more susceptible to chromosomal position effects (Figure 3.8 A-F). When  $\beta$ -galactosidase activity was observed in different tissues of independently transformed lines, a greater number were analysed to determine whether the staining patterns were related to *yp* expression. Interestingly, these effects did not occur as frequently with fragments that directed transcription from the native *yp3* promoter. This indicated that the *hsp70* promoter sequences in the pERI transformation vector were more susceptible to the influences of nearby enhancers than the *yp3* promoter, which could therefore contain negative regulatory elements or silencers involved in determining the overall pattern of *yp3* expression by preventing transcription in inappropriate tissues. A mechanism of repression is used to limit expression of the *Drosophila* homeotic genes, *Abdominal-B* and *Ultrabithorax*, to restricted domains along the embryo body axis (Müller and Bienz, 1991; Busturia and Bienz, 1993). Examples of other genes that have a temporal- and/or tissue-specific expression pattern regulated by silencers and enhancers acting in combination are the human  $\epsilon$ -globin gene (Wada-Kiyama et al., 1992), the rat collagen II gene (Savagner et al., 1990), the rat growth hormone gene (Larsen et al., 1986), the rat insulin I gene (Laimins et al., 1986), the chicken lysozyme gene (Baniahmad et al., 1987), the mouse heavy chain immunoglobulin gene (Kadesh et al., 1986) and the human  $\beta$ -interferon gene (Goodbourn et al., 1986). Interestingly, the *Drosophila* retrotransposon *gypsy* causes inactivation of upstream enhancers when it is inserted between these regulatory elements and a promoter, due to a

**Figure 3.8**

Photographs showing some interesting reporter gene expression patterns resulting from chromosomal position effects seen in transformed lines

- A** Germarium and single anterior/posterior follicle cells
- B** Previtelline space at the posterior of the oocyte
- C** Nerves innervating the ovary
- D** Cells at the tip of the male accessory gland
- E** Cells at the base of bristles seen in the thorax of males and females
- F** Muscle staining observed in males and females

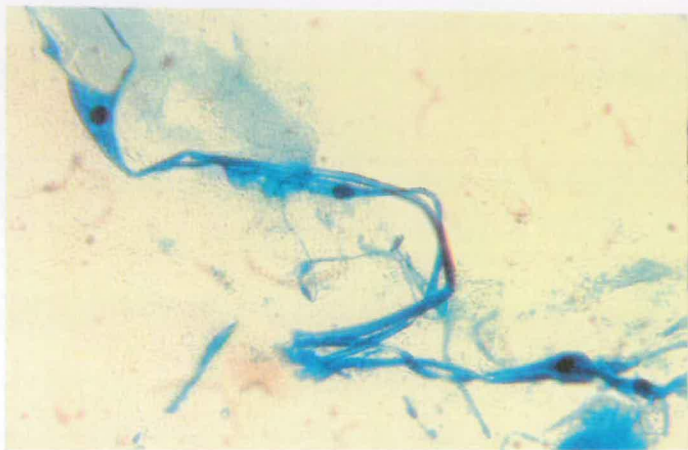
**A**                      **DISSECTED OVARY**



**B**                      **MATURE OOCYTE**



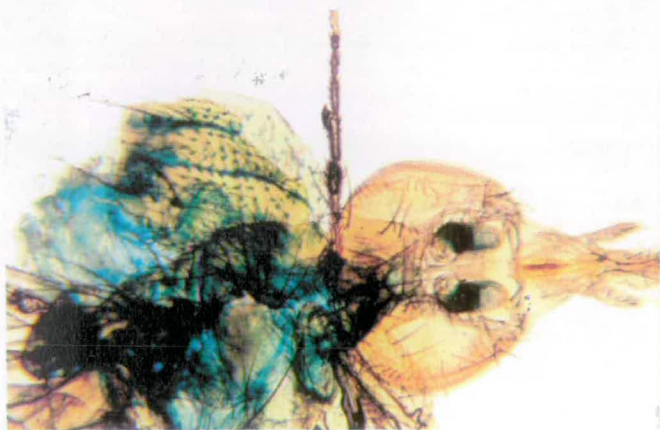
**C**                      **OVARIAN NERVOUS TISSUE**



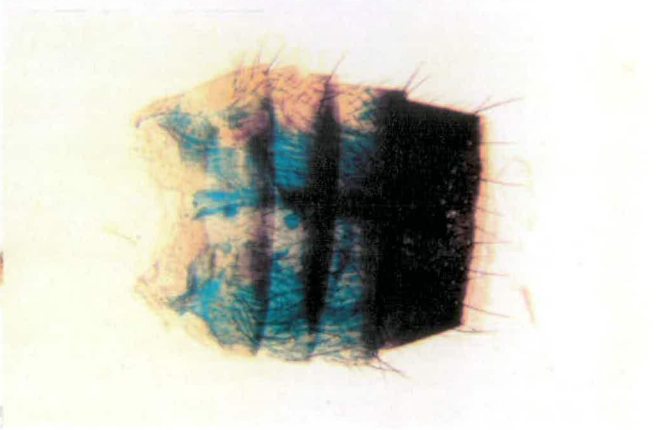
**D**                      **MALE TESTES**



**E**                      **MALE THORAX**



**F**                      **MALE ABDOMEN**



*suppressor of Hairy-wing* [*su(Hw)*] binding region within the *gypsy* element. The *su(Hw)* gene encodes a DNA-binding protein which, when bound to *gypsy*, interferes with interactions between enhancers and transcription factors required to activate the neighbouring gene (Geyer and Corces, 1992). This would be useful to incorporate into transformation vectors as it would isolate genes under study from the influences of chromosomal position effects (Roseman et al., 1993).

### **3.3.2 Sequence Comparisons Between the Identified Fat Body Enhancer Elements of *yp1*, *yp2* and *yp3***

Since the *yps* are coordinately transcribed and precisely regulated, it is likely that mechanisms and factors involved in their regulation will be the same for all three genes. We therefore carried out sequence comparisons between the identified fat body enhancer regions of *yp1*, *yp2* and *yp3* in order to identify conserved regulatory sequences. The 125 bp FBE located in the intergenic region separating *yp1* and *yp2* was compared to the 419 bp *yp3*FBE using the Bestfit program search, University of Wisconsin Genetics Computer Group, Sequence Analysis Software Package (Devereux et al., 1984), Version 7. The results of this are shown in Figure 3.9. The best sequence similarity occurred with the adjacent and overlapping AEF-1 and C/EBP binding sites shown by DNaseI footprinting assays to be present in the 125 bp FBE (Falb and Maniatis, 1992b). There was one gap in the *yp3* sequence, but a high level of 90 % similarity existed between the matched sequences (Figure 3.9B). Since BBF-2 was also reported to bind the 125 bp FBE (Abel et al., 1992) we searched for the recognition sequence in the *yp3*FBE and found that two putative sites existed, both with 78 % sequence similarity (Figure 3.9D) and one of which was a similar distance from the AEF-1 and C/EBP binding sites relative to the arrangement seen in the *yp1* and *yp2* FBE (Figure 3.9C i and ii). Binding studies are currently in progress to verify the significance of these findings. If these putative fat body factors do bind to the *yp3*FBE *in vitro*, footprinting assays, to confirm the sites of DNA contact, followed by mutational analysis of the binding

### Figure 3.9

A Part of the *yp1/yp2* FBE showing the location and target sequences of the AEF-1, C/EBP, BBF-2 and DSX binding sites

B Region of the *yp3*FBE with sequence similarity to the AEF-1 and C/EBP binding sites in the *yp1/yp2* FBE. The vertical lines indicate nucleotide identities, and the degree of similarity between the two sequences is also indicated.

Ci The AEF-1, C/EBP and BBF-2 binding sites identified in the *yp1/yp2* FBE by footprinting studies. Potential DSX binding sites are represented by ■

Cii The putative corresponding AEF-1, C/EBP and BBF-2 binding sites identified in the *yp3*FBE by computer sequence comparisons. The organisation of the binding sites can be seen to be similar with one of the BBF-2 recognition sequences separated from the overlapping AEF-1 and C/EBP binding sites by 44 bp in the *yp3*FBE compared with 52 bp in the *yp1/yp2* FBE. Potential DSX binding sites are represented by ■

D The computer sequence identities in the *yp3*FBE to the reverse complement sequence of BBF-2 binding site in the *yp1/yp2* FBE.

All nucleotide positions are given in relation to the transcription start site of the associated gene

A

AEF-1/CEBP

BBF-2

-317  
 AATCGTGCACAACCTACAATGTTGCAATCAGCGGAGCCTACAAAGTGATTACAAATTAAAAATATCAGGCGGCAGCAGGTGCTGCTAAGTCATCAGTGG  
 -217

DSX BINDING SITES

CTACAATGT

CTACAAAGT

TTACAAATT

CTGCTAAGT

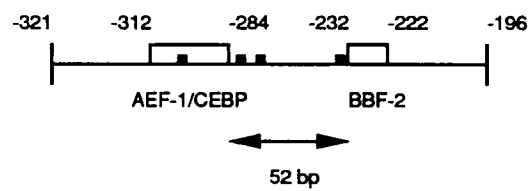
B

*yp1/2* FBE

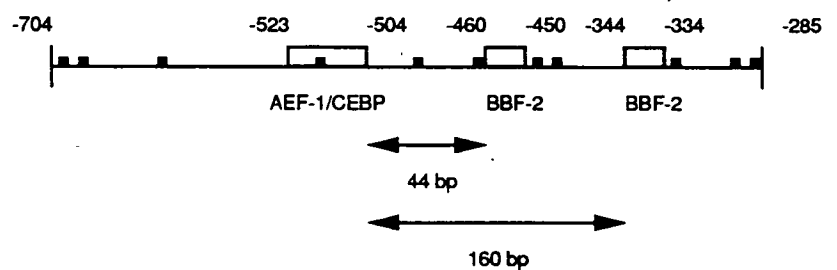
-311 TGCACAACCTACAATGTTGCAATCA -288  
 |||||  
 -523 TGCACAAGT . . . . TGTTCATCA -504  
 |||||  
 AEF-1 C/EBP

18/24nt (90%)

C i

*yp1/2* FBE

C ii

*yp3* FBE

D

*yp3**yp1/yp2* BBF-2

-460 CATTACTTTGG -450 -344 TAAATCTTAAC -334  
 |||||  
 GATGACTTAGC  
 6/11nt (78%)  
 7/11nt (78%)

sites and germ line transformation will reveal whether these factors are involved in regulating female fat body expression of *yp3* *in vitro*.

### 3.4 Summary and Discussion

Two *cis*-acting DNA regions have been identified and separated that are necessary for the ovarian and sex-specific fat body expression of the *Drosophila* *yp3* gene, both of which are located 5' of the *yp3* coding sequence. A 328 bp ovarian enhancer region (OE3) spanning nucleotides -285 to +43 was found to be sufficient to limit reporter gene expression to the correct follicle cell types at the expected stages 8-10 of oogenesis, reproducing the ovarian *yp* expression pattern. A neighbouring 419 bp sequence spanning nucleotides -704 to +43 was able to direct *lacZ* expression, from a heterologous promoter, in the fat body cells of adult females, but not males. Our experiments also demonstrated that these appeared to be the only regions essential for governing the observed *yp* expression pattern since the *yp3* coding, 3' and other 5' sequences did not contain any elements capable of independently directing the temporal-, tissue- or sex-specificity of expression. Nevertheless, it is possible that the identified *yp3*FBE or OE3 may contain more than one region capable of conferring tissue-specific expression, which could be determined by further dissecting these regions for analysis.

The sequence of *yp3* was compared to the *yp1/yp2* fat body enhancer sequence and the ovarian enhancers, OE1 and OE2, identified in transformation studies with the *yp1* and *yp2* genes (Garabedian et al., 1986; Logan et al., 1989; Logan and Wensink, 1990; Liddell and Bownes, 1991). The sequence similarities identified in *yp3* are presented in Figure 1.12 of the Introduction. Two *yp3* sequences were identified with sequence similarity to the *yp1/yp2* FBE. One was upstream of the *yp3* transcription start site between nucleotides -924 to -909, and the other was located downstream of the coding sequence between nucleotides +2081 to +2107. However, the results of transformation experiments with various *yp3* upstream, downstream and coding regions have shown that sequences for fat body-specific expression are located between nucleotides -704 to -285, suggesting that the



sequences identified by computer sequence comparisons are not necessary for the expression of *yp3* in the fat body. A comparison between OE1, OE2 and *yp3* yielded similar results. Two predicted sites of functional importance for the ovarian expression of *yp3* were identified in the upstream region between nucleotides -640 to -632, and between nucleotides -1577 to -1566. A third sequence similarity was identified in the *yp3* coding sequence between nucleotides +74 to +108. Transformation studies, however, localised sequences responsible for the ovarian transcription of *yp3* to a region spanning nucleotides -285 to +43, that does not contain any of the sequences identified in the computer sequence comparisons. However, it is possible that some of these regions are involved in the regulation of *yp* transcription by altering levels or rates of transcription.

It was not investigated in this study whether the identified *yp3*FBE and OE3 enhancers contained regulatory sequences that could interact with other regions of the *yp3* gene, but this should not be dismissed until they have been tested in conjunction with the *yp3* coding and downstream regions, or sub-fragments from these regions. It is possible that levels of transcription would be altered if the coding and 3' sequences were tested in conjunction with the identified enhancer elements, or if the 3' region was tested in a location downstream of the *lacZ* reporter gene. A variety of developmentally regulated genes expressed in a cell-type or tissue-specific manner are known to be influenced by downstream enhancers. The human *keratin 19* gene has an enhancer located approximately 1 kb from the 3' end of its transcript that is essential for its positive regulation in HeLa cells (Hu and Gudas, 1994). The human *keratin 1* gene is also regulated by a 3' region of 4.3 kb containing sequences that direct the cell-type specificity of expression and limit transcription to differentiating keratinocytes (Huff et al., 1993). Other examples of genes reported to contain 3' enhancers are the mouse cytokeratin *endo A* gene (Takemoto et al., 1991), the *immunoglobulin* genes (Meyer and Neuberger, 1989; Lauster et al., 1993; Matthias and Baltimore, 1993), the *globin* genes (Boddine and Ley, 1987; Kollias et al., 1987; Choi and Engel, 1988; Perez-Stable and Costantini, 1990), the  $\alpha$ - and  $\beta$ -chains of T-cell receptors (Krimpenfort et al., 1988; Winoto and Baltimore, 1989), the *Adh-1* gene of *Drosophila mulleri* (Fischer and Maniatis, 1988) and the steroid-

responsive *Drosophila Eip28/29* gene which contains ecdysone response elements in 3' flanking DNA sequences (Cherbas et al., 1991).

Computer sequence comparisons have identified potential binding sites, in the 419 bp *yp3FBE*, for the AEF-1, C/EBP and BBF-2 transcription factors reported to bind to the 125 bp FBE able to direct the sex-specific fat body expression of *yp1* and *yp2*. The experimental data demonstrating that the 419 bp fragment is sufficient to direct sex-specific fat body reporter gene expression, and the presence of potential binding sites for the AEF-1, C/EBP and BBF-2 transcription factors within this sequence, does suggest that these factors could be involved in regulating the sex- and tissue-specific expression of all three *yp* genes. If this is the case, the interactions which mediate this could also be of a similar nature. It is thought that AEF-1 and C/EBP compete for their overlapping binding sites in the 125 bp FBE since AEF-1 can displace pre-bound C/EBP from its recognition sequence (Falb and Maniatis, 1992b). This indicates that small changes in the cellular concentration of either protein could affect a sensitive on/off mechanism to determine whether the associated gene should be active or silent. Such a mechanism could be used to vary levels of *yp* expression in response to genetic, hormonal or environmental signals, and may be responsible for the observed relationship between nutritional intake and levels of *yp* transcription which subsequently affects the rate of egg production in females (Bownes et al., 1988a). However, as mentioned previously, it is known that the 125 bp FBE can be deleted from the *yp1/yp2* intergenic region without the loss of sex-specific fat body expression. Computer searches have identified other putative AEF-1, C/EBP and BBF-2 binding sites in the intergenic spacer, but not in the same proximal arrangement as found in the FBE (Bownes et al., 1993). This suggests that overlapping recognition sites may not be essential to activate the *yp* genes in the female fat body, and that the factors involved may also interact over longer DNA sequences.

An additional observation from the studies to identify tissue-specific enhancers associated with the *yp* genes was that the sex- and tissue-specificity of expression were never separated. When constructs were found to direct fat body-specific expression of either the *yps* themselves or reporter genes, this was also

restricted to the adult female. This may be explained by the presence of DSX binding sites that overlap with the sites for the regulatory proteins thought to be involved in governing the fat body expression of the *yp* genes. This occurs in the 125 bp *yp1/yp2* FBE, and computer searches have revealed that other putative DSX binding sites exist outside this FBE. The hypothesized interactions could therefore occur with other fragments from the *yp1/yp2* intergenic region capable of directing sex-specific expression in the female fat body (Abrahamsen et al., 1993; Bownes et al., 1993). The same interactions could also regulate *yp3* fat body transcription as several potential DSX binding sites exist in the identified *yp3*FBE (Figure 3.9C ii).

Unlike the regulation of *yp1* and *yp2* ovarian expression, it appears that the control of *yp3* transcription in the ovary is governed by a single ovarian enhancer contained within 328 bp of sequence located directly upstream of the *yp3* coding sequence. Since no regulatory proteins had been identified that were involved in regulating ovarian transcription of the *yps*, this fragment was used in preliminary experiments to determine whether putative *trans*-acting proteins were present in ovarian nuclear extracts. This is discussed in the following chapter.

## **CHAPTER 4**

### **Results**

#### **Gel Retardation Assays**

## 4.1 Introduction

Experiments to identify tissue-specific enhancers associated with *yp3* revealed that a 328 bp fragment immediately upstream of the *yp3* coding sequence, spanning nucleotides -185 to +43 was sufficient to confer the ovarian *yp* expression pattern on a *lacZ* reporter gene.  $\beta$ -galactosidase activity was detected only in the ovarian follicle cells during stages 8-10 of oogenesis, and this region was therefore termed OE3. A neighbouring 419 bp sequence spanning nucleotides -704 to -285 was sufficient to direct reporter gene expression from a heterologous promoter in the fat body of adult females, reproducing the sex-limited pattern of *yp* expression observed in this tissue, and was therefore termed the *yp3*FBE. These were the only enhancer regions identified that were autonomously involved in regulating *yp3* expression.

Similar studies with the divergently transcribed *yp1* and *yp2* genes had revealed that multiple regulatory regions were involved in the control of fat body expression (Garabedian et al., 1986; Abrahamsen et al., 1993) and ovarian expression (Logan et al., 1989; Logan and Wensink, 1990). The *yp1/yp2* fat body enhancer (FBE) was the first enhancer element to be identified that was sufficient to direct the sex-specific *yp* expression pattern in the fat body of adult females (Garabedian et al., 1986). It was subsequently used to demonstrate the binding of several *trans*-acting factors *in vitro*. These were box B-binding factor-2 (BBF-2; Abel et al., 1992), the CCAAT/enhancer-binding protein (C/EBP) and the adult enhancer factor-1 (AEF-1; Falb and Maniatis, 1992a, 1992b), which have been discussed in section 1.5.3 of the Introduction and section 3.1 of Chapter 3. However, no attempts were made to identify ovarian tissue-specific *trans*-acting factors that could be involved in regulating the ovarian expression pattern of *yp1* and *yp2*, possibly because no single *cis*-acting element was responsible for their expression in this tissue. Nevertheless, due to their highly regulated pattern of expression, the *cis*- and *trans*-acting factors involved in the control of *yp* transcription are expected to be similar, if not the same, for all three genes. As a first step towards identifying potential target sequences for ovarian *trans*-acting proteins the enhancer region, OE3, able to confer the correct

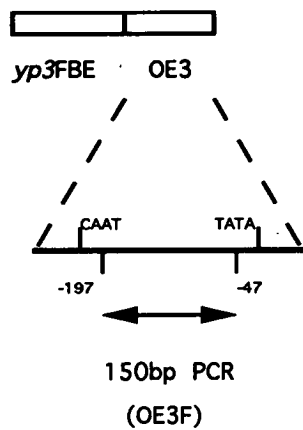
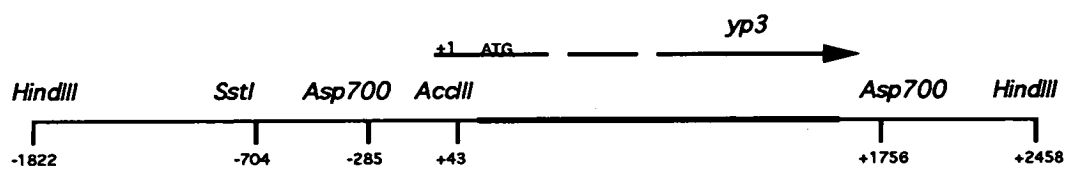
pattern of ovarian *yp3* transcription was used in gel retardation experiments to investigate whether proteins present in ovarian nuclear extracts could specifically bind to this construct.

## **4.2 Gel Retardation Assays with OE3**

Gel retardation assays were carried out with OE3 using *Drosophila* ovarian nuclear extracts prepared according to a previously reported method (Frank et al., 1992), with slight alterations (section 2.4.4, materials and methods). Since the 328 bp OE3 construct contained the *yp3* TATA and CAAT-box promoter elements (Figure 4.1), this was used in experiments to establish binding conditions as these sequences would be recognised by their associated binding proteins, expected to be present in ovarian nuclear extracts. In order to demonstrate the presence of potential novel DNA-binding proteins involved in ovarian *yp* regulation, a 150 bp sub-fragment from OE3, spanning nucleotides -197 to -47, was isolated using the polymerase chain reaction (PCR). The PCR product was termed OE3F and lacked the TATA and CAAT-box promoter elements (Figure 4.1). Preliminary gel-retardation experiments were performed with OE3 and OE3F to determine the levels of background carrier DNA (salmon testes DNA) that would prevent excess binding by non-specific proteins, but in the same instance, would allow binding of proteins that recognise specific target sequences within the probes. Examples of the results obtained with OE3 and OE3F are given in Figures 4.2A and 4.2B respectively. In the gel retardation assays with OE3, 20ng of <sup>32</sup>P end-labelled fragment and 3µg of crude ovarian nuclear extract were used in each reaction, and the amount of salmon testes DNA varied with 2µg, 1µg, 500ng, 250ng and 100ng in each reaction. The reactions were incubated on ice for 30 minutes and under these binding conditions, reducing the levels of carrier DNA resulted in increased protein binding, observed by reduced mobility of the probe through the gel. The optimum amount of carrier DNA was 500ng which was a 25-fold excess over the probe concentration. At this concentration, all of the labelled probe was bound and entered the gel (Figure 4.2A, lane 4). This probably included some non-specific binding since less protein was

#### Figure 4.1

*Drosophila yp3* gene illustrating the identified ovary (OE3) and fat body (*yp3*FBE) *cis*-regulatory regions. The 150 bp PCR sub-fragment (OE3F) derived from OE3 is also shown. OE3 and OE3F represent the ovarian constructs used in the gel retardation assays. Nucleotide positions of restriction enzyme sites and PCR boundaries are given relative to the *yp3* transcription start site.





**Figure 4.2A**

Autoradiograph of gel retardation assays with OE3. Each reaction contained approximately 20ng of <sup>32</sup>P end-labelled fragment. Lane 1 represents free (non-protein bound) DNA. Lanes 2-6 represent the DNA probe incubated with ovarian nuclear protein extract and decreasing amounts of salmon testes carrier DNA, the concentrations of which are given in the table:

Lane	Protein Concentration	Salmon Testes DNA (stDNA)	Molar Excess (stDNA:OE3)
1	—	1µg	—
2	3µg	2µg	100 X
3	3µg	1µg	50 X
4	3µg	500ng	25 X
5	3µg	250ng	12.5 X
6	3µg	100ng	5 X

**Figure 4.2B**

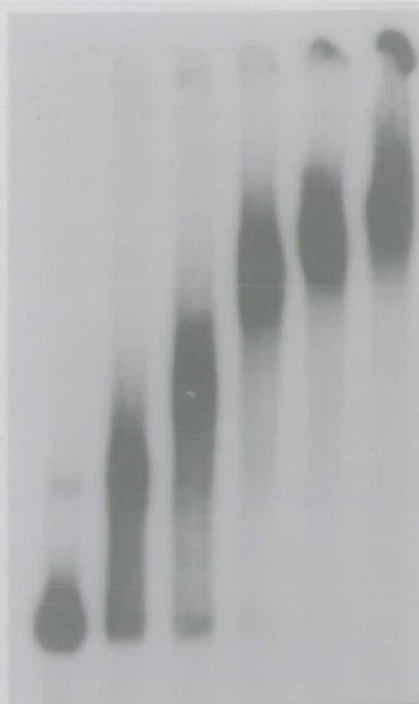
Autoradiograph of gel retardation assays with OE3F. Each reaction contained approximately 10ng of <sup>32</sup>P end-labelled fragment. Lane 1 represents free (non-protein bound) DNA. Lanes 2-5 represent the DNA probe incubated with ovarian nuclear protein extract and decreasing amounts of salmon testes carrier DNA, the concentrations of which are given in the table:

Lane	Protein Concentration	Salmon Testes DNA (stDNA)	Molar Excess (stDNA:OE3F)
1	—	1µg	—
2	3µg	1µg	100 X
3	3µg	500ng	50 X
4	3µg	250ng	25 X
5	3µg	100ng	10 X

A. OE3

1 2 3 4 5 6

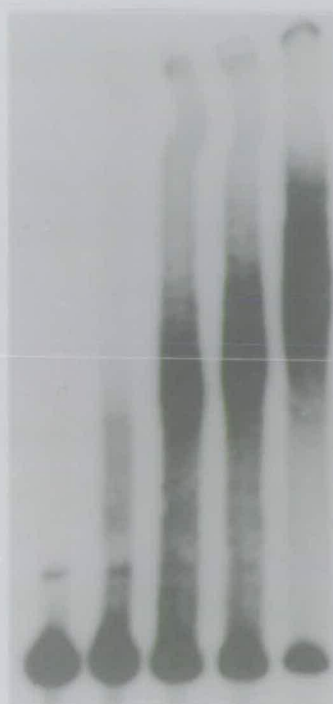
UNBOUND  
PROBE →



B. OE3F

1 2 3 4 5

UNBOUND  
PROBE →



bound with 2 $\mu$ g or 1 $\mu$ g of carrier DNA (Figure 4.2A, lanes 2 and 3 respectively). Although the mobility of the probe was still retarded in the presence of 2 $\mu$ g or 1 $\mu$ g of carrier DNA, less of the probe was bound than in the presence of 500ng of salmon testes DNA. Reducing the levels of carrier DNA to 250ng or 100ng (Figure 4.2A, lanes 5 and 6 respectively) resulted in more protein binding, but to the extent that not all of the probe could enter the gel. 500ng of salmon testes DNA was therefore concluded to be the optimum concentration of carrier DNA and was used in the next set of assays to determine the optimum level of non-specific competitor poly[d(I-C)].

In the preliminary gel-retardation assays with OE3F, 10ng of  $^{32}$ P end-labelled fragment and 3 $\mu$ g of crude ovarian nuclear extract were incubated with 1 $\mu$ g, 500ng, 250ng and 100ng of salmon testes carrier DNA in separate binding reactions. Similar results were obtained to those observed with OE3. A decrease in the level of carrier DNA resulted in an increase in the amount of protein binding, indicating that the binding complex involved interactions between the probe and several proteins, some of which were likely to be due to non-specific interactions. The optimum amount of carrier DNA in this set of reactions was 100ng, or a 10-fold molar excess over the probe concentration. Under these binding conditions, the maximum amount of OE3F was bound that could still enter the gel. 20ng of OE3F and 250ng of salmon testes DNA was therefore used in the next set of binding reactions to determine the optimum level of non-specific competitor poly[d(I-C)].

The addition of poly[d(I-C)] to the binding reaction eliminates non-specific binding of proteins to the probe. Results from gel-retardation assays with OE3 and OE3F, containing gradually increasing amounts of poly[d(I-C)] are given in Figures 4.3A and 4.3B respectively. Assays with OE3 contained 20ng of  $^{32}$ P end-labelled fragment, 3 $\mu$ g of crude ovarian nuclear extract, 500ng of salmon testes carrier DNA and increasing amounts (10ng, 20ng, 50ng, 100ng, 200ng, 400ng, 600ng, 800ng and 1 $\mu$ g) of non-specific competitor poly[d(I-C)]. The mobility of the DNA-protein complex increased on the addition of 100ng, or a 5-fold molar excess (Figure 4.3A, lane 6) and 400ng, or a 20-fold molar excess (Figure 4.3A, lane 8) of poly[d(I-C)], indicating the removal of non-specific proteins from the binding complex. Further addition of poly[d(I-C)] to 1 $\mu$ g, or a 50-fold molar excess (Figure

**Figure 4.3A and B**

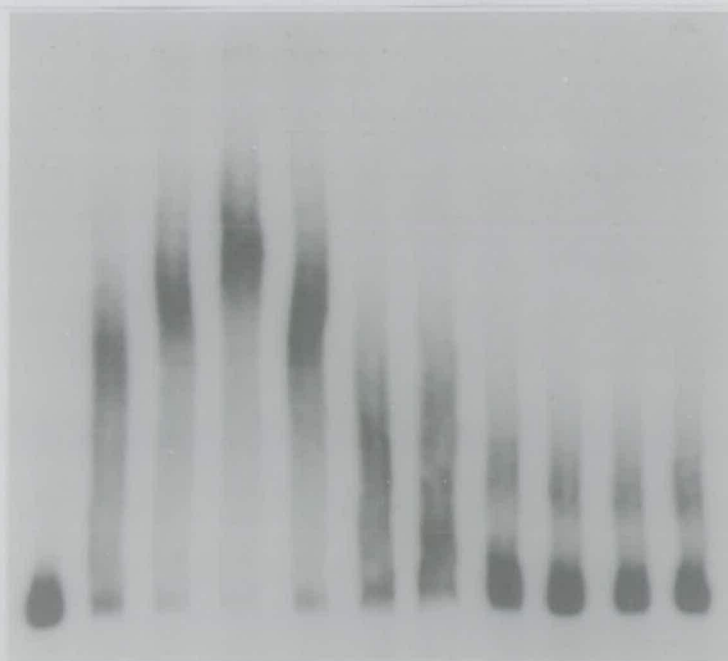
Autoradiographs of gel retardation assays with OE3 (A) and OE3F (B) that include poly[d(I-C)] in the binding reactions to eliminate binding by non-specific proteins. Each reaction contained approximately 20ng of end-labelled probe. In each autoradiograph, lane 1 represents free (non-protein bound) DNA. The remaining reactions each contained 3µg protein from ovarian nuclear extracts and lane 2 in each figure therefore represents the DNA-protein complex. The concentrations of salmon testes carrier DNA (stDNA) and poly[d(I-C)] used in each reaction are given in the table:

A. OE3				B. OE3F		
Lane	stDNA	poly[d(I-C)]	Molar Excess d(IC):OE3	stDNA	poly[d(I-C)]	Molar Excess d(IC):OE3F
1	500ng	—	—	250ng	—	—
2	"	—	—	"	—	—
3	"	10ng	0.5 X	"	10ng	0.5 X
4	"	20ng	1.0 X	"	20ng	1.0 X
5	"	50ng	2.5 X	"	50ng	2.5 X
6	"	100ng	5.0 X	"	100ng	5.0 X
7	"	200ng	10 X	"	200ng	10 X
8	"	400ng	20 X	"	400ng	20 X
9	"	600ng	30 X	"	600ng	30 X
10	"	800ng	40 X	"	800ng	40 X
11	"	1µg	50 X	"	1µg	50 X

A. OE3

1 2 3 4 5 6 7 8 9 10 11

UNBOUND  
PROBE →

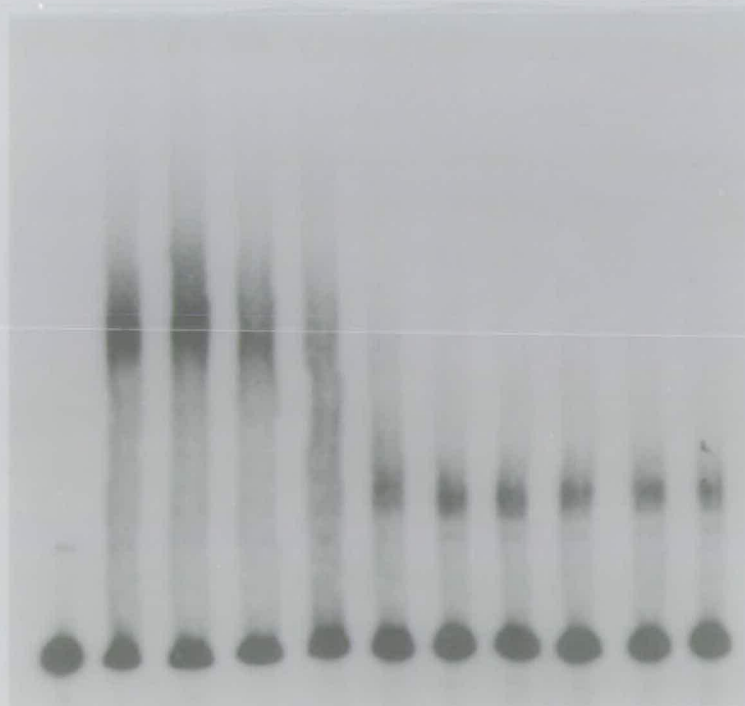


← B

B. OE3F

1 2 3 4 5 6 7 8 9 10 11

UNBOUND  
PROBE →



← B

4.3A, lane 11) did not further inhibit the formation of the protein-DNA complex, indicating that this resulted from specific interactions.

A similar set of reactions were performed using the OE3F PCR product (Figure 4.3B). The incubations contained 20ng of  $^{32}\text{P}$  end-labelled fragment, 3 $\mu\text{g}$  of crude ovarian nuclear extract, 250ng of salmon testes carrier DNA and increasing amounts (10ng, 20ng, 50ng, 100ng, 200ng, 400ng, 600ng, 800ng and 1 $\mu\text{g}$ ) of non-specific competitor poly[d(I-C)]. The mobility of the DNA-protein complex again increased on the addition of 100ng poly[d(I-C)] to the binding reaction (Figure 4.3B, lane 6), indicating that a non-specific protein component had been removed from the complex. Addition of poly[d(I-C)] to a 50-fold molar excess, however (Figure 4.3B, lane 11), did not abolish the formation of the remaining protein-DNA complex, suggesting that this could be due to specific interactions between OE3F and a protein or proteins present in the ovarian nuclear extract.

To further demonstrate the specificity of interactions between OE3 and OE3F with proteins present in ovarian nuclear extracts, gel-retardation assays were performed that included unlabelled OE3 or OE3F as specific competitor DNA, and the further addition of poly[d(I-C)] as non-specific competitor DNA. An example of the results obtained with OE3 is given in Figure 4.4A, and with OE3F in Figure 4.4B. Assays with OE3 contained 20ng of  $^{32}\text{P}$  end-labelled fragment, 3 $\mu\text{g}$  of crude ovarian nuclear extract, 500ng of salmon testes DNA, 500ng of poly[d(I-C)] and either increasing amounts (20ng, 50ng, 100ng, 200ng, 300ng) of unlabelled OE3 as specific competitor DNA, or increasing amounts (50ng, 100ng, 250ng, 500ng, 1 $\mu\text{g}$ ) of poly[d(I-C)] as non-specific competitor DNA. The binding complex was competed away from the end-labelled probe on the addition of 100ng, or a 5-fold molar excess of unlabelled OE3 to the binding reaction (Figure 4.4A, lane 5). However, the addition of a further microgram, or a 50-fold molar excess of poly[d(I-C)] to the binding reaction did not prevent the binding of ovarian nuclear proteins to target sequences within OE3 (Figure 4.4A, lane 12). This demonstrated that the components and conditions of the binding reactions were sufficient for protein-DNA interactions to occur, and that the observed interactions between OE3 and proteins present in ovarian nuclear extracts were specific.

# **Figure 4.4A and B**

Autoradiographs of gel retardation assays demonstrating the specificity of interactions between OE3 (A) and OE3F (B) with protein(s) present in ovarian nuclear extracts by specific competition with unlabelled probe, or non-specific competition with poly[d(I-C)]. Approximately 20ng of end-labelled probe was used in each reaction. In each autoradiograph, lane 1 represents free (non-protein bound) DNA. The remaining reactions each contained 3µg ovarian nuclear protein and lane 2 in each autoradiograph represents the DNA-protein complex. In the OE3 assays, 500ng of salmon testes carrier DNA and 500ng of poly[d(I-C)] were included in each reaction to eliminate non-specific protein binding. In the OE3F assays, 250ng of salmon testes carrier DNA and 250ng of poly[d(I-C)] were included in each reaction. The concentrations of unlabelled (cold) probe (for specific competition) and poly[d(I-C)] (for non-specific competition) used in each reaction are given in the table:

A. OE3				B. OE3F		
Lane	Cold OE3	poly[d(IC)]	Molar Excess	Cold OE3F	poly[d(IC)]	Molar Excess
1	—	—	—	—	—	—
2	—	—	—	—	—	—
3	20ng	—	1.0 X	20ng	—	1.0 X
4	50ng	—	2.5 X	50ng	—	2.5 X
5	100ng	—	5.0 X	100ng	—	5.0 X
6	200ng	—	10 X	200ng	—	10 X
7	300ng	—	15 X	300ng	—	15 X
8	—	50ng	2.5 X	—	50ng	2.5 X
9	—	100ng	5.0 X	—	100ng	5.0 X
10	—	250ng	12.5 X	—	250ng	12.5 X
11	—	500ng	25 X	—	500ng	25 X
12	—	1µg	50 X	—	1µg	50 X

A. OE3

1 2 3 4 5 6 7 8 9 10 11 12

UNBOUND  
PROBE →

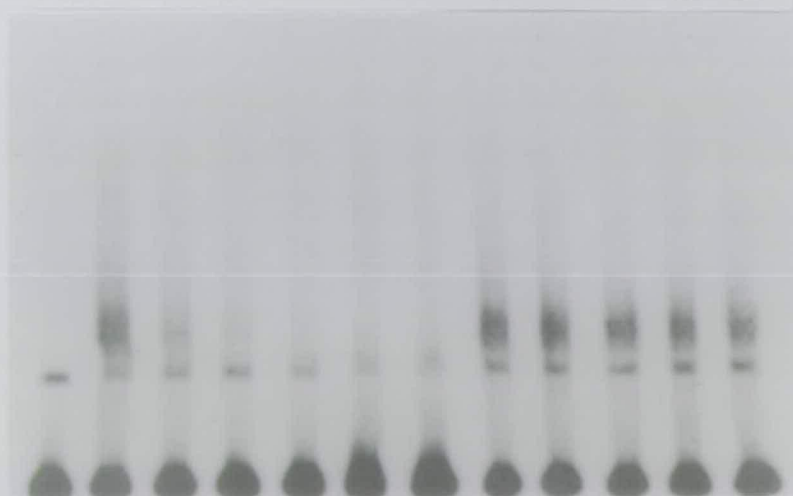
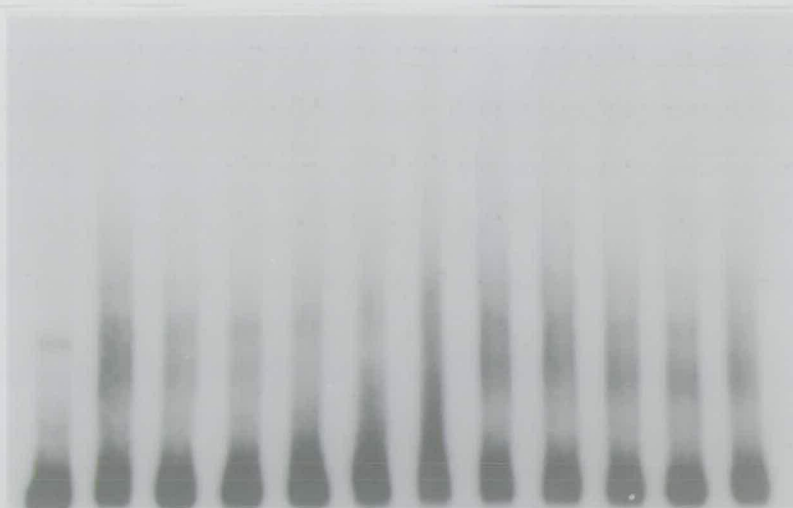
← B

B. OE3F

1 2 3 4 5 6 7 8 9 10 11 12

UNBOUND  
PROBE →

← B





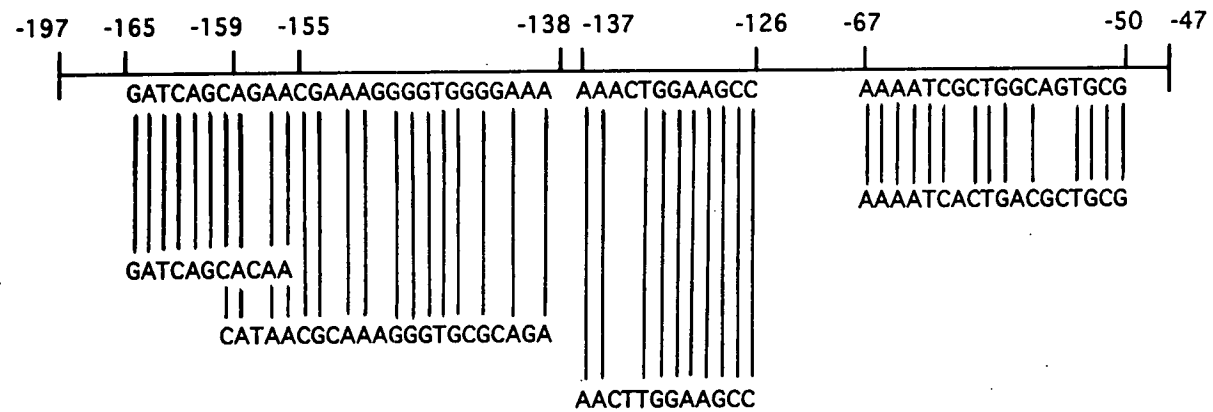
A similar set of reactions with OE3F were performed (Figure 4.4B). Binding reactions contained 20ng of  $^{32}\text{P}$  end-labelled fragment, 3 $\mu\text{g}$  of crude ovarian nuclear extract, 250ng of salmon testes DNA, 250ng of poly[d(I-C)] and increasing amounts of either unlabelled OE3F (20ng, 50ng, 100ng, 200ng, 300ng) as specific competitor DNA or poly[d(I-C)] (50ng, 100ng, 250ng, 500ng, 1 $\mu\text{g}$ ) as non-specific competitor DNA. 100ng, or a 5-fold molar excess of OE3F completely competed away the binding complex (Figure 4.4A, lane 5), whereas 1 $\mu\text{g}$ , or a 50-fold molar excess of poly[d(I-C)] did not affect complex formation (Figure 4.4A, lane 12). Therefore the binding that occurs between ovarian nuclear proteins and OE3F also appears to be the result of specific interactions.

#### **4.3 Sequence Comparisons Between the Identified Ovarian Enhancer Elements of *yp1*, *yp2* and *yp3***

In order to identify sequences that could be of functional importance for the activity of ovarian enhancer elements associated with the *yps*, computer searches were carried out between OE1, OE2 and OE3F using the Bestfit program search, University of Wisconsin Genetics Computer Group, Sequence Analysis Software Package, Version 7 (Devereux et al., 1984). Four putative sites of conserved sequence were identified (Figure 4.5). Two OE1 sequences of 11 bp and 18 bp had 91 % and 78 % identity to regions of OE3F respectively, and two OE2 sequences of 12 bp and 22 bp were matched to sequences in OE3F with 83 % and 73 % respective sequence similarity. Since all three *yps* have the same pattern of expression in the ovary, these could be candidate binding sites for ovarian tissue-specific transcription factors involved in regulating *yp* transcription in this tissue.

### Figure 4.5

Sequence similarities between OE1, OE2 and the OE3 PCR sequence, OE3F. Sequences from OE3F are given at the top of the figure and vertical lines indicate nucleotide identities to sequences from OE1 or OE2. The number of matched nucleotides and the percent identity between sequences are shown on the right-hand side of the figure. Nucleotide positions of OE1 and OE2 sequences are given relative to the transcription start site of *yp2* and the OE3F nucleotide positions relative to the *yp3* transcription start site



OE3 PCR sequence	(-197 to -47)
OE1 (-145 to -128)	14/18nt (78%)
OE1 (-106 to -116)	10/11nt (91%)
OE2 (+80 to +59)	16/22nt (73%)
OE2 (+36 to +47)	10/12nt (83%)

#### 4.4 Summary and Discussion

A single 328 bp ovarian enhancer element, OE3, spanning the *yp3* promoter region (nucleotides -285 to +43), had been identified that was sufficient to direct the ovarian pattern of *yp3* expression. In order to demonstrate binding of ovarian nuclear proteins to OE3, gel retardation assays were performed with *Drosophila* ovarian nuclear extracts. Since both the TATA and CAAT-box promoter elements were present, this allowed the conditions for nuclear protein binding to be established. The formation of protein-DNA complexes were observed (Figures 4.2A and 4.3A) one of which appeared to result from specific interactions (Figure 4.4A) likely to at least involve the binding of ovarian nuclear proteins to these promoter sequences.

Gel retardation assays were also carried out with a 150 bp sub-fragment from OE3, spanning nucleotides -197 to -47 (Figure 4.1), termed OE3F. Although this did not contain the TATA and CAAT-box promoter sequences, binding complexes were still observed (Figure 4.2B and 4.3B) one of which was not competed off by increasing concentrations of non-specific poly[d(I-C)] in the binding reactions, and therefore likely to be the result of specific protein-DNA interactions (Figure 4.4B). Computer searches between OE3F and previously identified binding sites for known transcription factors (such as AEF-1, C/EBP, BBF-2, EcRE, CF1/USP, OCT-1, SP-1) did not reveal any obvious sequence similarities (data not shown). It is therefore possible that the protein (or protein complex) that associates with OE3F is a novel tissue-specific factor involved in the control of ovarian *yp3* transcription.

Computer sequence comparisons between OE1, OE2 and OE3F have identified regions of sequence similarity which could represent putative binding sites for ovarian tissue-specific transcription factors. Although the results from computer searches alone cannot be interpreted as conclusive, they can be used to indicate sequences of functional importance when supported by experimental evidence. Since OE1 and OE2 are both required for the correct cell-type and developmental-specificity of ovarian *yp1/yp2* expression (Logan et al., 1989; Logan and Wensink,

1990), and OE3 contains regions of sequence similarity to both OE1 and OE2, it is possible that these sequences have a role in regulating ovarian *yp* transcription.

Future experiments to determine the significance of these findings would include P-element mediated germ-line transformation of OE3F, which could be tested upstream of the *hsp70-lacZ* reporter gene fusion in the pERI vector. Since regions of OE3 are deleted in the OE3F construct, one could expect that reporter gene expression would be observed in different subpopulations of follicle cells or during alternative stages of oogenesis, if the effects of OE3 result from a combination of positive and negative influences of OE3 regions on ovarian *yp3* transcription, as has been reported for OE1 (Logan and Wensink, 1991). It would also be of interest to carry out more gel retardation assays with OE3F to determine whether proteins in the binding complex could be competed for, firstly by OE1 or OE2, and secondly by the regions of OE1 and OE2 shown to have sequence similarity to OE3F. This would allow the significance of the results from computer sequence comparisons between OE1, OE2 and OE3F to be determined. In the long-term, the isolation of the ovarian DNA-binding protein(s) followed by DNaseI footprinting would identify the sites of protein-DNA interactions, and site-directed mutagenesis followed by P-element mediated germ-line transformation would determine whether these sequences were necessary for enhancer activity.

## **CHAPTER 5**

### **Results**

#### **The Effects of 20-hydroxyecdysone on *yp3* Expression**

## 5.1 Introduction

Regulation of *yolk protein (yp)* gene expression in *Drosophila melanogaster* is known to be complex, involving several factors that must all be present to function in a coordinated system that can activate or inactivate gene expression. There are three *yp* genes, and in order to maintain the regulated pattern of *yp* expression the somatic sex-determination gene hierarchy must be functional (reviewed by Wolfner, 1988; Baker, 1989; Belote, 1989; Slee and Bownes, 1990; Steinmann-Zwicky et al., 1990), the appropriate *cis*-acting tissue specific enhancers must be present (Garabedian et al., 1986; Logan et al., 1989; Logan and Wensink, 1990; Abrahamsen et al., 1993) and the nutritional environment of the fly must be adequate (Bownes and Blair, 1986; Bownes et al., 1988a). In addition, *yp* expression is known to be influenced by the insect hormones, juvenile hormone (Postlethwait and Handler, 1979; Jowett and Postlethwait, 1980; Bownes, 1986; Giorgi et al., 1993) and 20-hydroxyecdysone (Jowett and Postlethwait, 1980; Postlethwait et al., 1980; Bownes et al., 1983; Wu et al., 1987). 20-hydroxyecdysone mediates an effect on gene expression by interacting with ecdysone receptor molecules in association with the ultraspiracle protein. These complexes can subsequently recognise and bind to ecdysone response elements associated with their target genes (Yao et al., 1992; Thomas et al., 1993). Since injection of 20-hydroxyecdysone can induce *yp* expression in males (Postlethwait et al., 1980; Bownes et al., 1983), this provides an ideal system in which to identify the location of ecdysone response elements associated with the *yp* genes, as they are normally switched off in this sex. In order to identify response elements in females, flies have to be starved to reduce background levels of gene expression, which can remain variable. The effects of treatment with 20-hydroxyecdysone in males is therefore easier to monitor as there are no background levels of *yp* expression to consider. Also, the use of males in ecdysone induction experiments should not bias the outcome of the results as the aim is to identify the location of ecdysone response elements associated with the *yps*, which should be functional in both males and females.

### 5.1.1 The Effects of 20-hydroxyecdysone on *yp1/yp2* Gene Expression

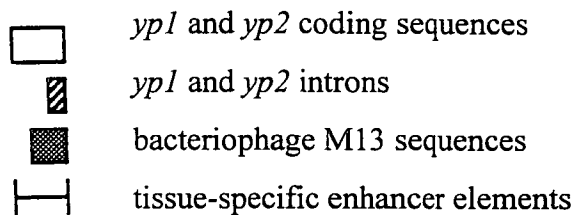
In order to identify the presence of 20-hydroxyecdysone responsive sequences associated with *yp1* and *yp2*, a series of DNA fragments spanning the *yp1* and *yp2* genomic region were tested in males (Figure 5.1). These constructs had been used in analyses to locate the *cis*-regulatory DNA sequences necessary for the fat body-specific pattern of *yp1* and *yp2* expression in adult females (discussed in sections 1.5.1 and 1.5.2 of the Introduction). As a result of these analyses a 125 bp fat body element (FBE) was located 196 bp upstream of the *yp1* transcription start site that was sufficient for the sex-specific fat body expression of both genes. All of the constructs tested for 20-hydroxyecdysone responsiveness contained the FBE as it may not have been possible to induce fat body expression in males in the absence of this tissue-specific enhancer.

The constructs tested for the presence of ecdysone response elements included a 4.8 kb genomic fragment containing the *yp1* and *yp2* coding sequences with the complete intergenic region (Figure 5.1A). In this construct the *yp* genes were tagged with bacteriophage M13 DNA inserted in the second exons such that transcripts could be distinguished from those of the endogenous *yps* (Garabedian et al., 1985). The construct illustrated in Figure 5.1B contained the complete *yp2* coding sequence, the intergenic region and a truncated *yp1* gene (Tamura et al., 1985). Following treatment of males with 20-hydroxyecdysone, levels of *yp2* transcripts or protein products could be assessed in a *yp2*<sup>-</sup> background, and truncated *yp1* transcripts would be distinguishable from those of the full length endogenous *yp1* gene in Northern blots. Three other constructs used in experiments to identify the FBE were tested for responsiveness to 20-hydroxyecdysone. One consisted of the *yp1* gene with 886 bp of upstream sequence, which was again tagged with M13 DNA sequences (Figure 5.1C). The other two constructs contained different regions from the intergenic region fused upstream of a *hsp70-lacZ* reporter gene (Garabedian et al., 1986). The regions spanned nucleotides -196 to -321 that encompassed the FBE (Figure 5.1D) and nucleotides -321 to -886 upstream of the *yp1* transcription start site (Figure 5.1E). An additional construct was examined for responsiveness to 20-hydroxyecdysone that was not used in experiments to locate the FBE. This consisted of a 929 bp *HindIII-AvaII* *yp1* fragment containing



Figure 5.1

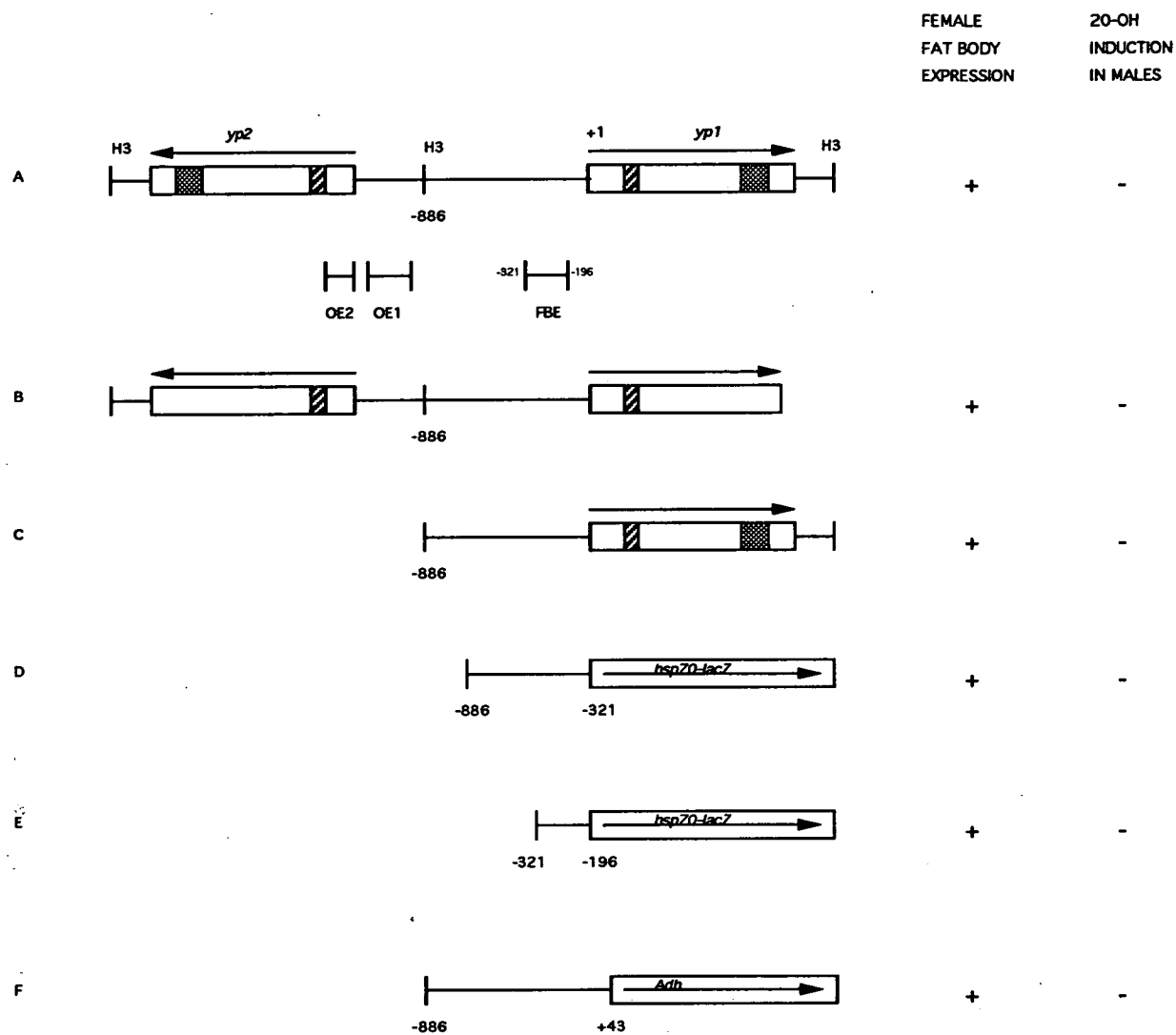
Constructs used in preliminary investigations to identify 20-hydroxyecdysone (20-OH) response elements in the intergenic region between the *yp1* and *yp2* genes of *Drosophila melanogaster*. The experiments are described in the text



Transcripts are represented as arrows with the arrowhead representing the direction of transcription. The alcohol dehydrogenase (*Adh*) and  $\beta$ -galactosidase (*hsp70-lacZ*) reporter genes are shown with the appropriate constructs. Nucleotide positions associated with the genomic fragments are given in relation to the *yp1* transcription start site (+1)

+ indicates expression

- indicates no expression



sequences spanning -886 to +43, with respect to the *yp1* transcription start site, fused to the *Drosophila Adh* gene (Figure 5.1F). Adh activity was detected only in females transformed with this construct, and the fat body was the only tissue that stained significantly (Shirras and Bownes, 1987).

Following treatment with 20-hydroxyecdysone, expression was not induced in males transformed with any of these constructs (Bownes and Shirras, unpublished results, constructs A, C, D and E; Postlethwait, personal communication, construct B; Shirras and Bownes, 1987, construct F). The only response to 20-hydroxyecdysone occurred with untransformed, wild type males, i.e. with the native *yp* genes, as demonstrated by the presence of yolk proteins on Western blots. This suggested that if any ecdysone response elements were associated with the *yp1* and *yp2* genes, they would be located outside the 4.8 kb genomic region studied.

To further examine 20-hydroxyecdysone induction of *yolk protein* gene expression we decided to assess the effects of 20-hydroxyecdysone on isolated regions of *yp3*. This gene had been isolated on a 4.3 kb genomic fragment and a series of reporter constructs containing upstream, coding, or downstream fragments, or a combination of these regions, fused upstream of a *lacZ* reporter gene had already been created, and transformed lines used to identify the location of tissue-specific *cis*-acting elements required to direct the expression of *yp3* in the fat body and ovarian follicle cells of adult females (Results Chapter 3). A reporter construct containing the isolated *yp3* downstream region would be of particular interest in ecdysone induction experiments since equivalent regions downstream of *yp1* and *yp2* had not been tested on their own. A response to 20-hydroxyecdysone had not been observed in males transformed with any of the *yp1* and *yp2* constructs investigated, that had included the coding regions and/or upstream sequences. It was therefore thought that if any ecdysone response elements were present, they were likely to be located in sequences downstream of these divergently transcribed genes since these were the only regions lacking in the constructs tested.

### 5.1.2 The Effects of 20-hydroxyecdysone on *yp3* Gene Expression

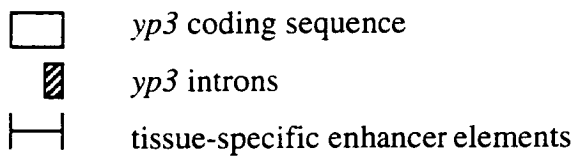
The effects of 20-hydroxyecdysone on *yp3* expression were investigated using males transformed with various *yp3* constructs. Ten constructs were examined in total (Figure 5.2, constructs A-J). Construct A represents the 4.3 kb genomic fragment on which the *yp3* gene was isolated. This contained the *yp3* coding sequence, 825 bp of 3' flanking DNA and 1822 bp of 5' flanking sequence. It had previously been demonstrated that a response to 20-hydroxyecdysone occurred in males transformed with this construct (D.Mauchline, personal communication) which suggested that ecdysone response elements had been isolated with the *yp3* gene. Construct B differed from construct A, having only 704 bp of 5' flanking DNA. However, following treatment with 20-hydroxyecdysone, *yp3* gene expression was also induced in *yp3*<sup>-</sup> males transformed with this construct. With both these constructs, induction of gene expression was monitored by detection of the protein product. YP3 was detected on protein gels by Western blots or <sup>35</sup>S labelling followed by autoradiography (D.Mauchline, personal communication).

### 5.2 *yp3* Constructs Tested for Responsiveness to 20-hydroxyecdysone

Constructs C to J had been used in experiments to identify the location of tissue-specific *cis*-acting elements required to direct the observed pattern of *yp3* expression in the fat body and ovarian follicle cells of adult females. Constructs C,D and E retained the *yp3* promoter region and were therefore placed directly upstream of the *lacZ* reporter gene contained within the pCaSpeR-AUG-βgal transformation vector (Thummel et al., 1988). Constructs F to J that lacked the *yp3* promoter region were cloned into a modified version of this transformation vector (pERI) that contained sufficient *Drosophila hsp70* promoter sequences to allow these fragments to be tested for their effects on reporter gene activity (Ingolia et al., 1980; Dudler and Travers, 1984). Construct C was a 1865 bp fragment containing the region upstream of *yp3* spanning nucleotides -1822 to +43, with respect to the *yp3* transcription start site, and construct D consisted of a shorter 747 bp upstream fragment spanning nucleotides -704

Figure 5.2

The *yp3* constructs tested for the presence of 20-hydroxyecdysone (20-OH) response elements. The 4.3 kb *yp3* genomic fragment is illustrated at the top of the diagram with the restriction enzyme sites used to create the reporter constructs C-J illustrated below. The nucleotide positions of these restriction sites are given with in relation to the *yp3* transcription start site (+1)

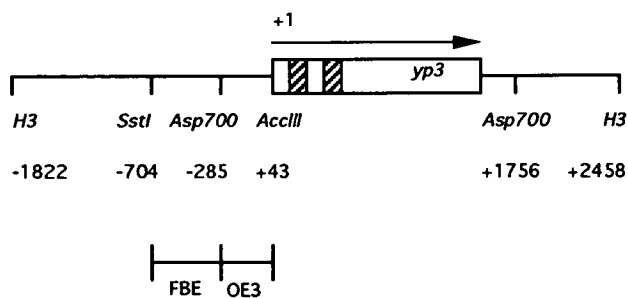


Transcripts are represented as arrows with the arrowhead representing the direction of transcription. Constructs C, D and E that contained *yp3* promoter sequences were fused upstream of the *lacZ* reporter gene contained in the pCaSpeR-AUG- $\beta$ -gal transformation vector (Thummel et al., 1988). Constructs F-J that lacked *yp3* promoter sequences were cloned into this vector modified to contain a *hsp70-lacZ* gene fusion.

+ indicates expression

-indicates no expression

+/- indicates expression sometimes



FEMALE  
FAT BODY  
STAINING

20-OH  
INDUCTION  
IN MALES

A		+	+
B		+	+
C		+	+
D		+	+
E		-	+
F		+	+
G		+	+
H		-	+
I		-	+/-
J		-	-

to +43. Both of these constructs were sufficient to direct *lacZ* reporter gene expression in the fat body and ovaries of transformed females, demonstrating that sufficient promoter sequences were present for gene activation. Constructs E and F divided this 747 bp region into upstream fragments of 328 bp and 419 bp respectively. The 328 bp fragment contained sequences between nucleotides -285 and +43, which had been shown to be sufficient to direct reporter gene expression in the ovarian follicle cells of adult females during stages 8-10 of oogenesis, and was subsequently termed ovarian enhancer 3, or OE3. Likewise, it had been demonstrated that 419 bp fragment spanning nucleotides -704 to -285, was able to confer sex-specific reporter gene expression in female fat bodies from a heterologous promoter, and this region was therefore termed the *yp3* fat body enhancer or *yp3*FBE. This was the first construct to be tested upstream of the *hsp70-lacZ* reporter gene fusion, and demonstrated that the pERI transformation vector contained sufficient *hsp70* promoter sequences to activate gene expression *in vivo* (see section 3.2.2; Figure 3.5) (Ronaldson and Bownes, in press). The final upstream region to be tested for responsiveness to 20-hydroxyecdysone consisted of the *yp3*FBE and the remaining upstream sequences, namely a 1537 bp fragment spanning nucleotides -1822 to -285. To determine whether the *yp3* coding region and downstream sequences contained ecdysone response elements these regions were investigated together (construct H, a 2415 bp fragment composed of nucleotides +43 to +2458), and separately (Constructs I and J), by placing them upstream of the *hsp70-lacZ* reporter gene. Construct I was a 1713 bp fragment containing the *yp3* coding region on its own, from nucleotide positions +43 to +1756 which included only 150 bp of downstream sequence, and construct J was a 702 bp fragment containing only 3' flanking sequences spanning nucleotides +1756 to +2458.

This series of constructs effectively divided the *yp3* genomic DNA, known to respond to treatment with 20-hydroxyecdysone, into its upstream, coding and downstream components. All of the constructs were placed upstream of the *lacZ* reporter gene such that a response to treatment with 20-hydroxyecdysone in transformed males would be apparent by histological staining. It would also be possible to assess the responses to 20-hydroxyecdysone quantitatively as  $\beta$ -galactosidase activity could be monitored by spectrophotometry. From this analysis, it was hoped that the location of

ecdysone response elements associated with the regulation of *yp3* gene expression could be determined.

Both histochemical staining and spectrophotometric assays were therefore used to detect reporter gene expression, following treatment with 20-hydroxyecdysone, in males transformed with the *yp3* reporter constructs. For each experiment, at least twelve males from three independently transformed lines, were injected with 0.1 $\mu$ l of 10<sup>-2</sup> M 20-hydroxyecdysone and allowed to recover either overnight at 18°C, or for 6 hours at 25°C. These conditions were based on previous experiments determining the time course of induction of gene expression in males following an injection of 20-hydroxyecdysone (Bownes, 1982). Two males were then dissected and tested for gene activity, which could be detected by *lacZ* staining. The remaining ten males were dissected to separate the abdominal fat body from the carcass or thoracic fat body tissue. The abdominal fat bodies were pooled and the thoracic fat bodies were pooled to make separate extracts for the  $\beta$ -galactosidase assays. Extracts were prepared according to section 2.4.3.11 in the materials and methods. Colour reactions, representing  $\beta$ -galactosidase activity, were measured at 420nm on a Pye Unicam PU 8800 UV/VIS Spectrophotometer. Readings were taken at 10 minute intervals over a period of 2 hours. Prior to making extracts for the  $\beta$ -galactosidase assays, haemolymph was taken from the ten males and pooled for analysis on SDS-PAGE gels to ensure that the native *yps* were induced following treatment with 20-hydroxyecdysone. Proteins were visualised using the ISS Pro-Blue staining system (Integrated Separation Systems, Enprotech), an example of which is shown in Figure 5.3. The endogenous *yps* were consistently expressed at similar levels in males following treatment with 20-hydroxyecdysone, which demonstrated that the hormone was active. Any variations in reporter gene activity would therefore not be due to a lack of hormonal influence, but could result from the differences between reporter constructs. Also, once extracts had been made for the  $\beta$ -galactosidase assays, 50  $\mu$ l of the sample was retained and the levels of protein concentration determined according to the Bio-Rad Protein Assay. There was no significant difference between the protein yields of extracts from males transformed with the *yp3* constructs tested (data not shown). This demonstrated that any variations in  $\beta$ -galactosidase activity would not



### Figure 5.3

ISS Pro-Blue stained 10 % acrylamide linear protein gel. Haemolymph samples were taken from wild type Oregon R ( $Or^R$ ) males and females, and following treatment with 20-hydroxyecdysone, from males transformed with the pERI vector (V) and males transformed with constructs C, D, E, F and J (Figure 5.2). The  $Or^R$  male lane represents a pooled sample from 20 individuals, the  $Or^R$  female lane a pooled sample from 10 flies, and the haemolymph samples from transformed males were pooled from 30 flies. The position of the YPs are indicated

Or<sup>R</sup> ♂    Or<sup>R</sup> ♀    V    C    D    E    F    J



YP1  
YP2  
YP3

result from varying yields of protein between different extracts but could be due to different levels of reporter gene expression. The remainder of these extracts were subsequently used in spectrophotometric assays to determine the levels of  $\beta$ -galactosidase activity, that should reflect the levels of reporter gene expression.

As a control, following treatment with 20-hydroxyecdysone, levels of reporter gene expression were assessed in males transformed with the pERI transformation vector that contained the *hsp70-lacZ* reporter construct without any *yp* sequences. The readings obtained in these experiments are included in the results tables and demonstrated that background levels of reporter gene activity were low. This provided a baseline for comparison of the levels of  $\beta$ -galactosidase activity in experiments with the reporter constructs containing the *yp3* sequences. In this experimental situation, variations in  $\beta$ -galactosidase activity from the various *yp3* reporter constructs should reflect whether ecdysone response elements associated with *yp3* are present.

### **5.3 Results and Conclusions**

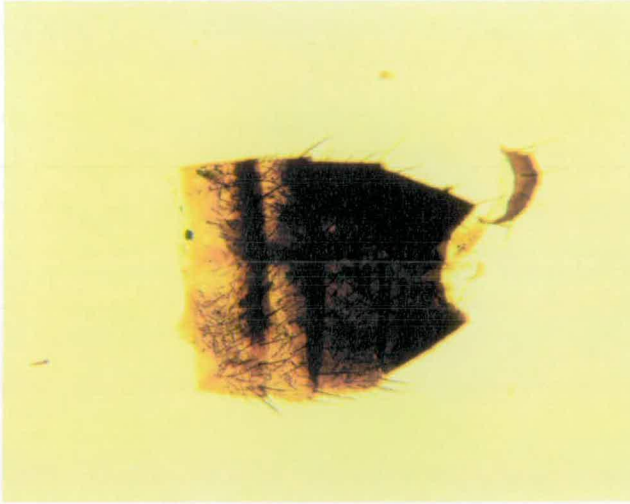
#### **5.3.1 Effects of 20-hydroxyecdysone on *lacZ* Expression in Males Transformed with Reporter Constructs Containing *yp3* Upstream Sequences**

The first constructs to be tested were C and D, containing 1865 bp (-1822 to +43) and 747 bp (-704 to +43) of *yp3* upstream sequence respectively. *LacZ* staining indicated that reporter gene expression was being induced in both the abdominal and thoracic fat body tissue of males. An example of the staining seen with construct C is shown in the photographs in Figure 5.4. Interestingly, the staining was consistently more intense in males transformed with construct D (the shorter construct of 747 bp composed of the *yp3* FBE and OE3), suggesting the possibility that a repressor binding site (or sites) preventing ecdysone induction had been removed that was contained within the upstream sequence spanning -704 bp to -1822 bp. This was verified on comparison of the levels of  $\beta$ -galactosidase activity in assays using males transformed with constructs C and D. Levels of reporter gene activity were assessed in both the

#### Figure 5.4

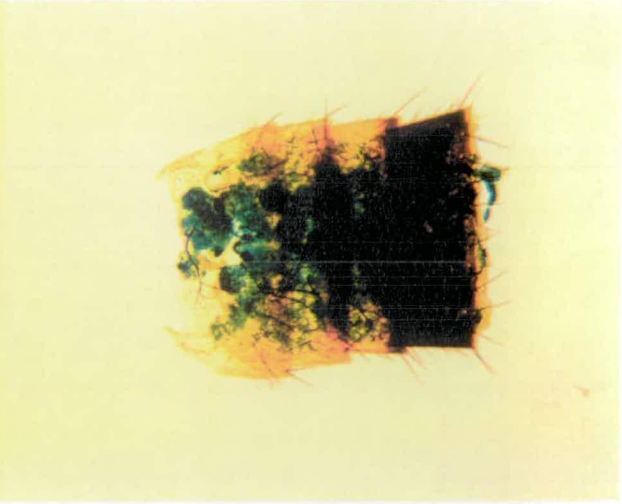
Photographs of the abdominal and thoracic fat body staining in males transformed with the *yp3* upstream region spanning -1822 to +43 fused to a *lacZ* reporter gene. The construct is represented below the photographs. No staining is observed in the abdomen or thorax of males that have not been treated with 20-hydroxyecdysone (photographs **A** and **C**). Blue staining representing *lacZ* expression and  $\beta$ -galactosidase activity is present in both the abdominal and thoracic fat body of males following treatment with 20-hydroxyecdysone (photographs **B** and **D**)

A



MALE ABDOMEN – 20-HYDROXYECDYSONE

B



MALE ABDOMEN + 20-HYDROXYECDYSONE

C

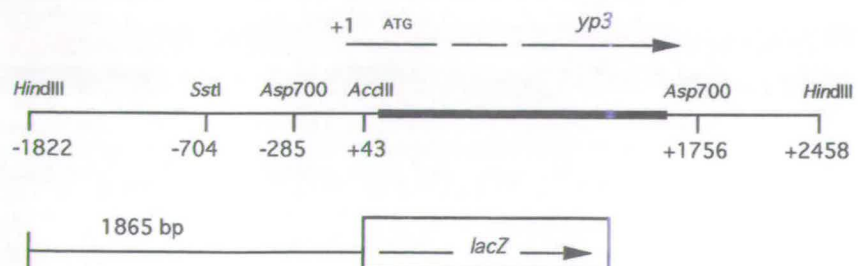


MALE THORAX – 20-HYDROXYECDYSONE

D



MALE THORAX + 20-HYDROXYECDYSONE



abdominal and thoracic fat body tissue as staining was observed in both these regions. The readings taken at 10 minute time intervals over a period of 2 hours are presented in table 5.1. The figures in bold are the final readings taken, which are used to compare the extent of reporter gene expression and to draw comparisons between the response to 20-hydroxyecdysone by different reporter constructs. These readings resulted from an average of three independently transformed lines for construct C and four independently transformed lines for construct D. For each line tested, ten males were injected with 20-hydroxyecdysone and allowed to recover before being dissected and pooled to make the extracts used in the assays. In order to determine the significance of the differences in reporter gene expression, or *lacZ* activity, between constructs, the readings in small print above and below represent the range from the lowest to the highest readings respectively, recorded for individual lines. It appears that the level of induction of reporter gene expression is approximately two-fold greater in males transformed with construct D compared to males transformed with construct C, and that this occurred in both the abdominal and thoracic fat body tissue. However, the lowest readings obtained for construct D overlapped with the highest readings obtained for construct C, and this occurred in both tissue locations. Although a two fold difference was consistently obtained on repetition of this experiment, it remained difficult to determine whether this was a significant difference.

To demonstrate that the observed induction was not male-specific and would occur with the native *yp* genes in females, the induction of reporter gene expression was also measured in the abdominal fat body of females transformed with construct D, as this upstream region of *yp3* produced the greatest response in the abdominal fat body of males. In order to reduce background levels of reporter gene expression that resulted from the influence of the sex-determination gene hierarchy and nutrition in females, flies were starved on 3 % agar for 48 hours prior to the application of 20-hydroxyecdysone. They were then allowed to recover (without food) at 18°C overnight before proceeding with the assays. Levels of reporter gene activity were also measured in starved females that were not treated with 20-hydroxyecdysone, to obtain levels of background reporter gene expression. The results of  $\beta$ -galactosidase assays with extracts prepared from the female abdominal fat bodies are shown in table 5.2 and the

### Table 5.1

Levels of  $\beta$ -galactosidase in the abdominal and thoracic fat body of males following treatment with 20-hydroxyecdysone. A colour reaction was measured spectrophotometrically in males transformed with the pERI vector (V), that contained no *yp3* sequences, males transformed with 1865 bp of *yp3* upstream sequence fused to a *lacZ* reporter gene (-1822 to +43, Figure 5.2 C) and males transformed with 747 bp of *yp3* upstream sequence fused to a *lacZ* reporter gene (-704 to +43, Figure 5.2 D). The readings taken at 10 minute time intervals are given and represent an average taken from three independently transformed lines for construct C and four independently transformed lines for construct D. The final figures in bold are the readings used to compare levels of reporter gene expression between constructs. The figures in small print, above and below the figures in bold, represent the lowest and highest readings obtained respectively, i.e. the range in levels of reporter gene activity that occurred between the three independently transformed lines

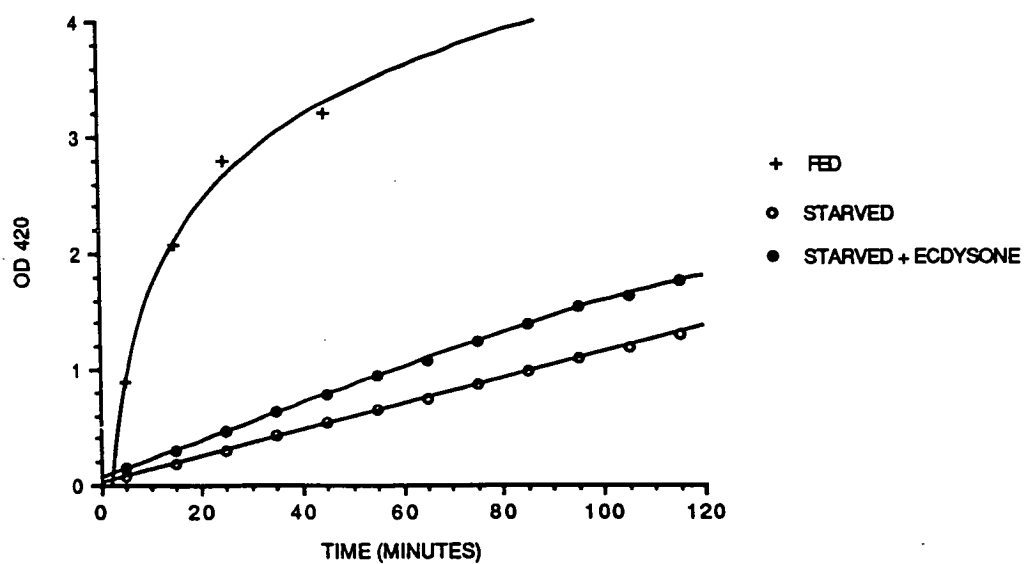
TIME (MINS)	Abdominal fat body			Thoracic fat body		
	V	C	D	V	C	D
	OD <sub>420</sub>	OD <sub>420</sub>	OD <sub>420</sub>	OD <sub>420</sub>	OD <sub>420</sub>	OD <sub>420</sub>
5	0.016	0.034	0.029	0.061	0.056	0.067
15	0.017	0.032	0.036	0.066	0.059	0.073
25	0.021	0.038	0.044	0.074	0.063	0.084
35	0.025	0.039	0.051	0.083	0.071	0.096
45	0.028	0.040	0.058	0.090	0.075	0.108
55	0.030	0.046	0.066	0.098	0.077	0.120
65	0.029	0.046	0.072	0.106	0.083	0.135
75	0.029	0.048	0.077	0.113	0.088	0.148
85	0.031	0.050	0.085	0.121	0.096	0.161
95	0.034	0.051	0.093	0.129	0.100	0.177
105	0.034	0.053	0.098	0.135	0.104	0.189
	0.03	0.05	0.07	0.14	0.09	0.12
115	0.036	<b>0.055</b>	<b>0.104</b>	0.143	<b>0.110</b>	<b>0.202</b>
	0.05	0.06	0.16	0.15	0.13	0.32



## Table 5.2

Levels of  $\beta$ -galactosidase activity in the abdominal fat body of females transformed with 747 bp of *yp3* upstream sequence fused to a *lacZ* reporter gene (-704 to +43, Figure 5.2 D), demonstrating that the response to 20-hydroxyecdysone is equivalent to that observed in the abdominal fat body of males. The results are averaged from three independently transformed lines and are also presented as a graph. In order to detect this response, females were starved to reduce background levels of reporter gene expression

TIME (MINUTES)	FED	STARVED	STARVED + ECDYSONE
5	0.895	0.078	0.140
15	2.083	0.188	0.295
25	2.789	0.303	0.459
35	-	0.417	0.623
45	3.201	0.532	0.777
55	-	0.647	0.945
65	-	0.745	1.075
75	-	0.865	1.238
85	-	0.976	1.386
95	-	1.096	1.535
105	-	1.192	1.636
115	-	1.289	1.756



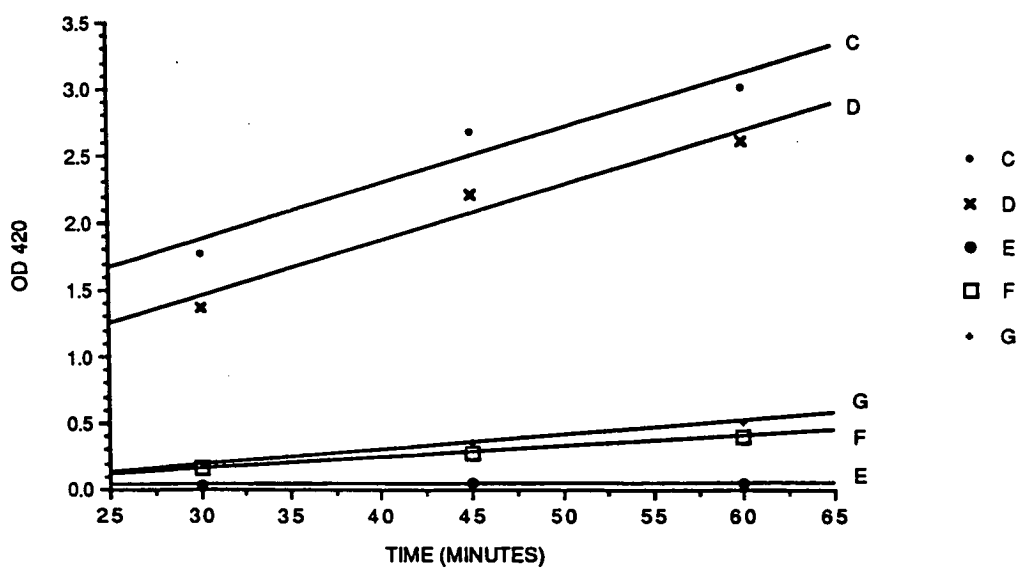
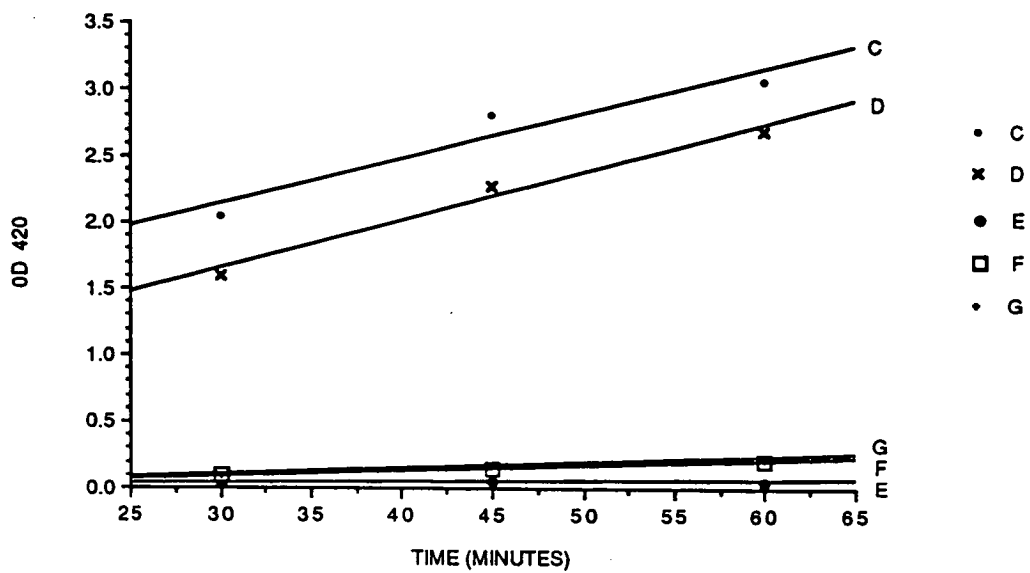
associated graph. A slight increase in *lacZ* expression occurred in starved females treated with hormone in comparison to females that did not receive an injection of 20-hydroxyecdysone, demonstrating that induction of reporter gene expression also took place in females transformed with construct D. This result provided evidence supporting the observed response in males and the conclusion that *cis*-acting hormone response elements were associated with the *yp3* gene.

An additional conclusion drawn from the experiments with males transformed with constructs C and D was that an element of repression could be involved in governing the overall effects of 20-hydroxyecdysone on *yp* expression, since the reporter gene fused to 704 bp of upstream sequence (construct D) was expressed more strongly than that fused to 1822 bp of *yp3* upstream sequence (construct C). A series of constructs that further dissected regions C and D were therefore used to investigate this in more detail. The effects of 20-hydroxyecdysone on reporter gene expression in males transformed with constructs E, F and G were investigated. Constructs E and F effectively divided the 704 bp construct D into two components of 328 bp (nucleotides +43 to -285) and 419 bp (nucleotides -285 to -704) containing either the ovarian or fat body tissue-specific enhancers respectively. Construct G was composed of the *yp3* FBE together with the 5' sequences, spanning nucleotides -285 to -1822) that included the region (-704 to -1822) that appeared to reduce or repress levels of induction of gene expression. In addition, it was of interest to investigate the *yp3* 3' flanking sequences isolated in construct J (nucleotides +1756 to +2458) since no induction of gene expression had been observed in any of the *yp1* and *yp2* constructs tested, which lacked the corresponding 3' region. The induction of *lacZ* expression using males transformed with constructs E, F, G and J was measured in both the abdominal and thoracic fat body tissue. Constructs C and D are included for direct comparison. It was apparent that most of the 5' constructs tested varied in response to 20-hydroxyecdysone, relative to each other, depending on the region of the fat body tissue used in the assays. The two sets of results from the abdominal and thoracic fat body tissue are therefore interpreted separately. The results are given in table 5.3 (section 5.3.1.1) for the abdominal fat body and table 5.4 (section 5.3.1.2) for the thoracic fat body, and are also presented as graphs. To ensure that the differences

Figure 5.5

**A** Graph representing the levels of  $\beta$ -galactosidase activity in the abdominal fat body of females transformed with constructs C, D, E, F and G

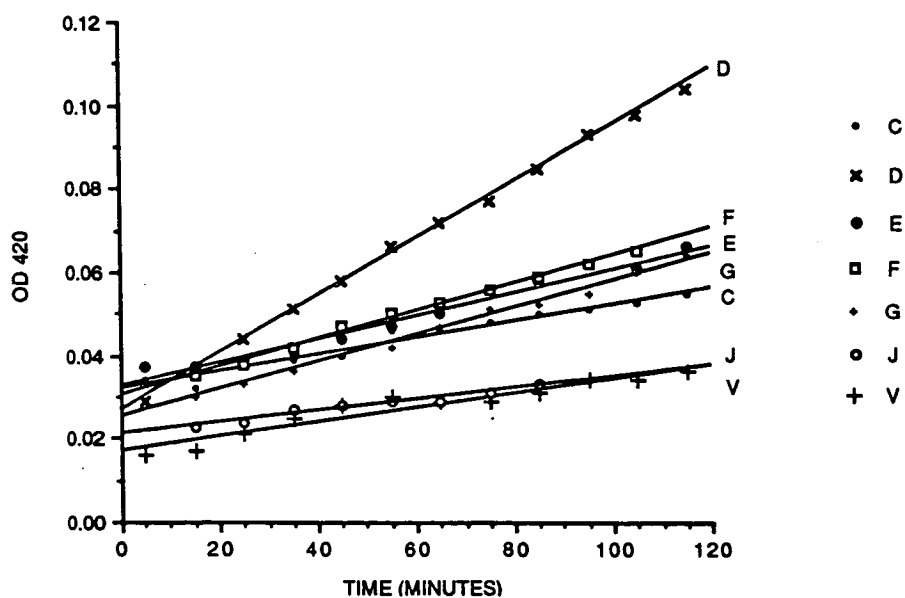
**B** Graph representing the levels of  $\beta$ -galactosidase activity in the thoracic fat body of females transformed with constructs C, D, E, F and G



### Table 5.3

Spectrophotometry readings representing levels of  $\beta$ -galactosidase activity, following 20-hydroxyecdysone treatment, in the abdominal fat body of males transformed with the pERI transformation vector (V) containing no *yp3* sequences, and males transformed with constructs C, D, E, F, G and J illustrated in Figure 5.2. The readings taken at 10 minute time intervals are given and represent an average taken from 3 independently transformed lines for constructs V, C, E, F and J, and 4 independently transformed lines for D and G. The figures in bold are the readings used to compare levels of reporter gene expression between constructs. The figures in small print, above and below the figures in bold, represent the lowest and highest readings obtained respectively, or the range in levels of reporter gene activity that occurred between the transformed lines. Levels of  $\beta$ -galactosidase activity are also presented in the form of a graph

TIME (MINS)	V OD <sub>420</sub>	C OD <sub>420</sub>	D OD <sub>420</sub>	E OD <sub>420</sub>	F OD <sub>420</sub>	G OD <sub>420</sub>	J OD <sub>420</sub>
5	0.016	0.034	0.029	0.037	-	0.029	-
15	0.017	0.032	0.036	0.037	0.035	0.030	0.023
25	0.021	0.038	0.044	0.039	0.038	0.033	0.024
35	0.025	0.039	0.051	0.041	0.042	0.036	0.027
45	0.028	0.040	0.058	0.044	0.047	0.040	0.028
55	0.030	0.046	0.066	0.047	0.050	0.042	0.029
65	0.029	0.046	0.072	0.050	0.053	0.047	0.029
75	0.029	0.048	0.077	-	0.056	0.051	0.031
							0.02
85	0.031	0.050	0.085	0.058	0.059	0.052	<b>0.033</b>
							0.05
95	0.034	0.051	0.093	-	0.062	0.055	-
					0.05		
105	0.034	0.053	0.098	0.061	<b>0.065</b>	0.060	-
					0.10		
	0.03	0.05	0.07	0.03		0.04	
115	<b>0.036</b>	<b>0.055</b>	<b>0.104</b>	<b>0.066</b>	-	<b>0.064</b>	-
	0.05	0.06	0.16	0.12		0.09	



between reporter constructs in the two fat body tissue locations were not due to the presence or absence of tissue-specific elements, and were in fact a response to 20-hydroxyecdysone, levels of reporter gene expression were measured in the abdominal and thoracic fat body tissue of females transformed with constructs C, D, E, F and G. The results of this are represented as graphs in Figure 5.5A and B and demonstrate that there is little difference between levels of reporter gene expression in females with different constructs relative to each other, whether the assays were carried out using abdominal or thoracic fat body tissue.

#### **5.3.1.1 Response to 20-hydroxyecdysone by *yp3* Reporter Constructs in the Abdominal Fat Body Tissue**

Levels of reporter gene expression in the abdominal fat body of males transformed with constructs C, D, E, F, G and J were assayed by measuring levels of  $\beta$ -galactosidase activity spectrophotometrically. The results are presented in table 5.3 and the associated graph. The greatest response to 20-hydroxyecdysone in the abdominal fat body tissue, shown to be statistically significant, occurred in males transformed with construct D, that contained both the *yp3* fat body and ovarian tissue-specific elements. The ovarian and fat body enhancers were separated in constructs E and F respectively, and when these were assayed separately they gave similar levels of reporter gene expression, but at reduced levels to the induction obtained when they were present together in construct D. It was difficult to determine if the reduced levels of reporter gene expression between constructs D and E, or D and F were significant but nevertheless, this result was reproducible and there are several possible explanations that could account for this. Fat body expression was induced in males in the absence of the *yp3* tissue-specific FBE, which suggests that the ovarian enhancer could contain a weak tissue-specific fat body element that is normally dormant, but could induce fat body expression in the presence of high levels of 20-hydroxyecdysone, and in the absence of other regulatory 5' sequences. Since the hormonal response is greater when both the ovarian and fat body elements are present (in comparison to when they are assayed separately) it is not unreasonable to suggest that a hormone response element in



the ovarian enhancer could be interacting with the *yp3*FBE to produce this result. It is also possible that OE3 could contain a 20-hydroxyecdysone response element that is able to function independently of the *yp3* FBE. If the fat body was the only tissue containing the appropriate transcription factors to interact with hormone response elements a response would only be seen in this tissue, regardless of the *cis*-acting tissue-specific enhancers present in reporter constructs. Another explanation for this observation is that a 20-hydroxyecdysone response element could be located at the position where OE3 and the *yp3*FBE are separated.

As mentioned earlier, comparison of *yp3* expression in males transformed with constructs C and D, showed that induction of reporter gene expression was lower in the abdominal fat body of males transformed with the upstream region spanning nucleotides -1822 to +43 (construct C), than males transformed with the shorter upstream fragment spanning nucleotides -704 to +43 (construct D) suggesting that a repressor binding site could be present in the upstream sequence spanning nucleotides -704 to -1822 upstream of the *yp3* transcription start site. This was investigated further by comparing the levels of reporter gene induction in males transformed with constructs F and G. Construct F contained the *yp3* FBE and construct G was composed of the FBE and the upstream region that contained the potential repressor binding site(s). Unlike the difference seen with constructs C and D, constructs F and G gave similar levels of reporter gene expression as reflected by levels of  $\beta$ -gal activity. However, this may have been due to the absence of the ovarian enhancer in constructs F and G. This would be consistent with the observed results obtained when the fat body and ovarian enhancer regions were tested together and individually. Induced levels of reporter gene expression were greater with the construct containing both these elements, than constructs containing either the ovary or fat body enhancers. In other words reporter gene expression was increased when the FBE was tested in the presence of the ovarian enhancer. To see if further evidence could be obtained to support the observation that the presence of the ovarian enhancer in *yp3* reporter constructs increased the response to 20-hydroxyecdysone, levels of induced reporter gene expression were compared between constructs C and G. Construct C contained all the *yp3* upstream sequences and construct G contained all the upstream sequences

except for OE3. However these two constructs produced a similar response to 20-hydroxyecdysone. Therefore, the presence or absence of the ovarian enhancer did not affect the response to 20-hydroxyecdysone when the upstream sequences spanning nucleotides -704 to -1822 were present. It therefore seems that a response to 20-hydroxyecdysone is only enhanced by OE3 in the absence of the -704 bp to -1822 bp upstream region, and -704 bp to -1822 bp region can only suppress the response to 20-hydroxyecdysone when the *yp3* promoter region is present. From these observations it seems that the *yp3* promoter region, or ovarian enhancer, and the upstream sequences spanning -704 to -1822 could be interacting with each other to determine the response to 20-hydroxyecdysone.

In conclusion, the *yp3* upstream sequences have been shown to contain 20-hydroxyecdysone response elements, and there is evidence to indicate that sequences within OE3 are having a positive influence on the response of reporter constructs to 20-hydroxyecdysone, and that there is at least one binding site for a repressor factor that can suppress the response to 20-hydroxyecdysone in the *yp3* upstream sequences between -704 bp to -1822 bp.

Finally, there was no response to this hormone in males transformed with the 3' sequences contained in construct J.

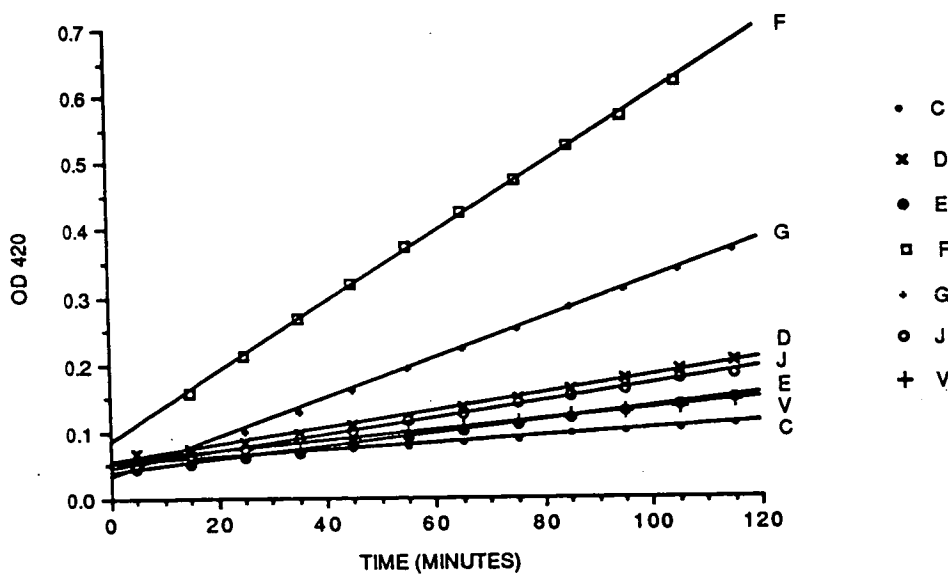
#### **5.3.1.2 Response to 20-hydroxyecdysone by *yp3* Reporter Constructs in the Thoracic Fat Body Tissue**

Most of the *yp3* reporter constructs containing 5' sequences assessed for responsiveness to 20-hydroxyecdysone in the abdominal fat body of transformed males had a different response, with respect to each other, in the thoracic fat body of the same males. Results from the  $\beta$ -galactosidase assays with the thoracic fat body tissue are presented in table 5.4 and the associated graph. The only comparable responses in the two fat body regions were observed with constructs C and D. As in the abdominal fat body, levels of  $\beta$ -galactosidase activity were approximately two-fold greater in males transformed with construct D (containing 747 bp of *yp3* upstream sequence spanning

Table 5.4

Spectrophotometry readings representing levels of  $\beta$ -galactosidase activity, following treatment with 20-hydroxyecdysone, in the thoracic fat body of males transformed with the pERI transformation vector (V) with no *yp3* sequences, and males transformed with constructs C, D, E, F, G and J illustrated in Figure 5.2. The readings taken at 10 minute time intervals are given and represent an average taken from 2 independently transformed lines for constructs V and F, 3 independently transformed lines for constructs C, E, G and J, and 4 independently transformed lines for construct D. The figures in bold are the readings used to compare levels of reporter gene expression between constructs. The figures in small print, above and below the figures in bold, represent the lowest and highest readings obtained respectively i.e. the range that occurred between the independently transformed lines. The results are also presented in the form of a graph

TIME (MINS)	V OD <sub>420</sub>	C OD <sub>420</sub>	D OD <sub>420</sub>	E OD <sub>420</sub>	F OD <sub>420</sub>	G OD <sub>420</sub>	J OD <sub>420</sub>
5	0.061	0.056	0.067	0.045	-	0.057	0.052
15	0.066	0.059	0.073	0.051	0.156	0.075	0.063
25	0.074	0.063	0.084	0.060	0.211	0.100	0.075
35	0.083	0.071	0.096	0.069	0.265	0.130	0.090
45	0.090	0.075	0.108	0.080	0.317	0.160	0.101
55	0.098	0.077	0.120	0.090	0.372	0.194	0.115
65	0.106	0.083	0.135	0.099	0.423	0.222	0.126
75	0.113	0.088	0.148	0.108	0.472	0.251	0.140
85	0.121	0.096	0.161	0.119	0.524	0.281	0.151
95	0.129	0.100	0.177	0.128	0.569	0.309	0.162
105	0.135	0.104	0.189	0.137	0.50 <b>0.621</b> 0.75	0.338	0.176
115	0.14 <b>0.143</b> 0.15	0.09 <b>0.110</b> 0.13	0.12 <b>0.202</b> 0.32	0.13 <b>0.147</b> 0.17	-	0.27 <b>0.366</b> 0.48	0.17 <b>0.184</b> 0.22



nucleotides -704 to +43) compared to males transformed with construct C (containing 1865 bp of *yp3* upstream sequence spanning nucleotides -1822 to +43).

Unlike the results obtained in  $\beta$ -galactosidase assays with the abdominal fat body extracts, a similar response was monitored in the thoracic fat body of males transformed with constructs F and G. Construct F contained the *yp3*FBE and construct G contained the *yp3*FBE and the upstream region spanning nucleotides -704 to -1822, containing potential repressor binding sites. Induction of reporter gene expression in males transformed with construct F was approximately two-fold that observed in males transformed with construct G.

These results provided supportive evidence for the presence of a repressor binding site in the *yp3* upstream region spanning nucleotides -704 to -1822. Unlike the results obtained with the abdominal fat body, these results also indicated that the repressive effect of this upstream region could occur in the presence or absence of OE3. In the abdominal fat body, a response to 20-hydroxyecdysone was only reduced by the region spanning -704 to -1822 when OE3 was also present. This indicated that, in the thoracic fat body, sequences in this upstream region were not interacting with the *yp3* promoter region to suppress the response to 20-hydroxyecdysone, but could do so independently. An explanation for this difference could be that different tissue-specific factors are present in the fat body in the abdominal and thoracic regions, such that different interactions are occurring between the *yp3* upstream response elements and the available *trans*-acting factors.

To further investigate the theory that different interactions could be occurring between *yp3* upstream sequences and factors present in the abdominal or thoracic fat body tissue, the response to 20-hydroxyecdysone in the thoracic fat body of males transformed with construct D (a 747 bp fragment composed of the *yp3*FBE and OE3) was compared to that in males transformed with construct F (the *yp3*FBE on its own). In the abdominal fat body the response to 20-hydroxyecdysone was greater in males transformed with construct D, than males transformed with construct F. In the thoracic fat body however, levels of reporter gene expression and hence  $\beta$ -galactosidase activity, were greater in males transformed with construct F than males transformed with construct D, suggesting that, in the thoracic fat body tissue, OE3 was having a

negative effect on the induction of gene expression by 20-hydroxyecdysone. This was confirmed on comparison of the levels of  $\beta$ -galactosidase activity obtained in males transformed with construct C, containing the entire *yp3* upstream region from nucleotide position -1822 to +43, and males transformed with construct G, containing the upstream region spanning nucleotides -1822 to -285 that lacked OE3. A greater response to 20-hydroxyecdysone occurred in males transformed with construct G compared to males transformed with construct C.

In the thoracic fat body, it therefore appeared that the *yp3* promoter region, or OE3, was repressing the levels of reporter gene induction by 20-hydroxyecdysone. This was also an independent effect since this occurred in the presence or absence of the upstream region spanning nucleotides -704 to -1822. Unlike the effects observed in the abdominal fat body, it seems that this upstream region and OE3 act independently to influence the response to 20-hydroxyecdysone in the thoracic fat body.

As in the abdominal fat body, there was no response to 20-hydroxyecdysone in the thoracic fat body of males transformed with the *yp3* 3' sequences contained in construct J.

### **5.3.2 Effects of 20-hydroxyecdysone on *lacZ* Expression in Males Transformed with Constructs Containing *yp3* Coding and Downstream Sequences**

To date, no experiments have been carried out on the *yp* genes with 20-hydroxyecdysone to determine whether any ecdysone response elements are present in the coding sequence itself, in which case the gene could be involved in its own regulation. Construct I contained the *yp3* coding region (with only 150 bp of 3' untranslated sequence and 16 bp of non-coding sequence) placed upstream of the *lacZ* reporter gene and this was tested in males. Following treatment with 20-hydroxyecdysone, slight staining was observed in the abdominal fat body, which again demonstrated that the 125 bp tissue-specific FBE was not required to produce a response in the fat body. Interestingly, when this region was tested in conjunction with the 3' sequences, using males transformed with construct H, the staining that reflected

the extent of induced reporter gene expression, was much greater (D.Mauchline and M.Bownes, personal communication). Photographs showing the response with constructs H and I are given in Figure 5.6. This suggested that both sequences downstream of *yp3*, from nucleotides +1756 to +2458, and the coding sequence of *yp3* itself contained hormone response elements that could interact with each other to mediate a response to 20-hydroxyecdysone, since no expression occurred when the 3' region was tested on its own, and the response when the coding region was tested alone was also much weaker. Alternatively, a 20-hydroxyecdysone response element could be present at the site separating the *yp3* coding and downstream sequences.

### **5.3.3 Computer Searches to Locate Ecdysone Response elements in the 4.3 kb *yp3* Genomic Sequence**

Experimental evidence indicated that 20-hydroxyecdysone response elements were present throughout the 4.3 kb *yp3* genomic region studied. Consistent with these results was the outcome of computer searches that we carried out with consensus sequences for ecdysone response elements. This revealed possible 20-hydroxyecdysone response elements in the *yp3* downstream region, in the coding sequence between the two introns, and in OE3, the sequence immediately upstream of the transcription start site able to direct the ovarian-specific transcription of *yp3*. The site in the 3' region of *yp3* had sequence similarity to a 23 bp hyphenated dyad sequence identified in the *Drosophila hsp27* promoter that was shown to confer ecdysone inducibility on a heterologous gene. The sequence was GACAAGGGTTCAATGCACTTGTC, where the underlined bases highlight the hyphenated dyad symmetry (Riddiough and Pelham, 1987). The region of *yp3* with sequence similarity to this spanned nucleotides +2377 to +2398 in the 3' flanking region, and is shown in Figure 5.7A. The sequence similarity between the two introns was derived from the ultraspiracle binding site located upstream of the *Drosophila sl5* chorion gene, with the sequence AGGTCACGT (Shea et al., 1990; Christianson et al., 1992). This shared the sequence motif GGTC A reported to be a possible common ancestor DNA motif for invertebrate and vertebrate hormone response elements

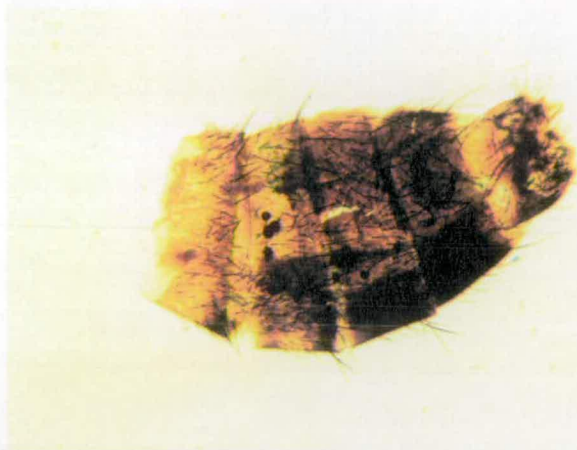
### Figure 5.6

**A and B**      Photographs representing the response to 20-hydroxyecdysone in the abdominal fat body of males transformed with the construct containing the *yp3* coding region fused upstream of a *hsp70-lacZ* reporter gene. The construct is represented below the photographs. Reporter gene expression is not induced by treatment with 20-hydroxyecdysone, although faint staining is sometimes observed

**C and D**      Photographs representing the response to 20-hydroxyecdysone in the abdominal fat body of males transformed with the construct shown below containing the *yp3* coding and 3' sequences fused upstream of a *hsp70-lacZ* reporter gene. The blue staining demonstrates that this construct was responsive to treatment with 20-hydroxyecdysone

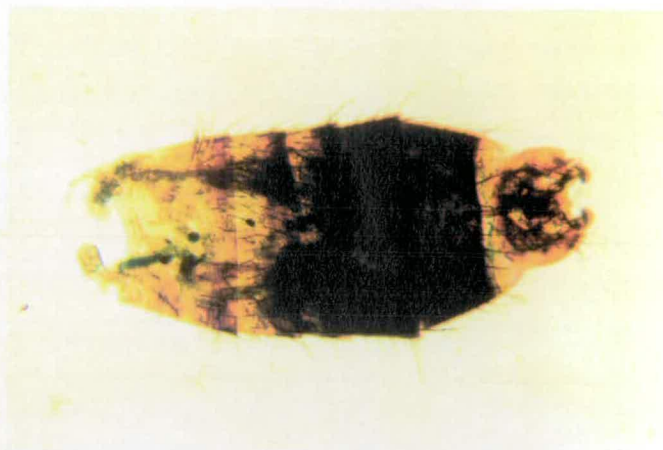


A

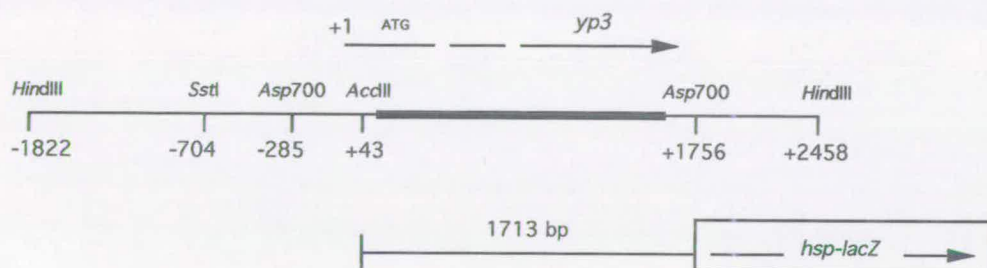


MALE ABDOMEN - 20-HYDROXYECDYSONE

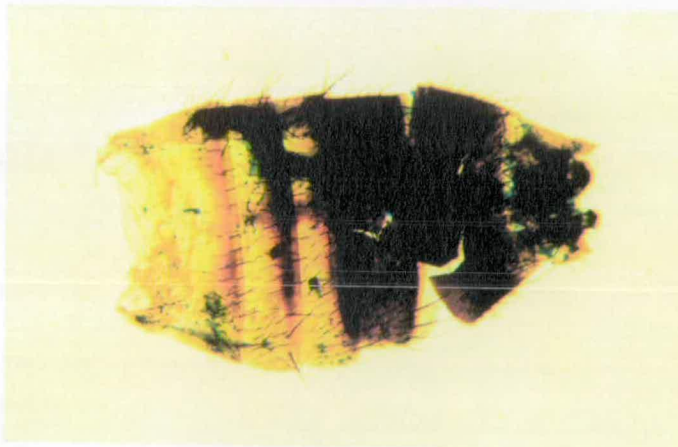
B



MALE ABDOMEN + 20-HYDROXYECDYSONE



C



MALE ABDOMEN - 20-HYDROXYECDYSONE

D



MALE ABDOMEN + 20-HYDROXYECDYSONE

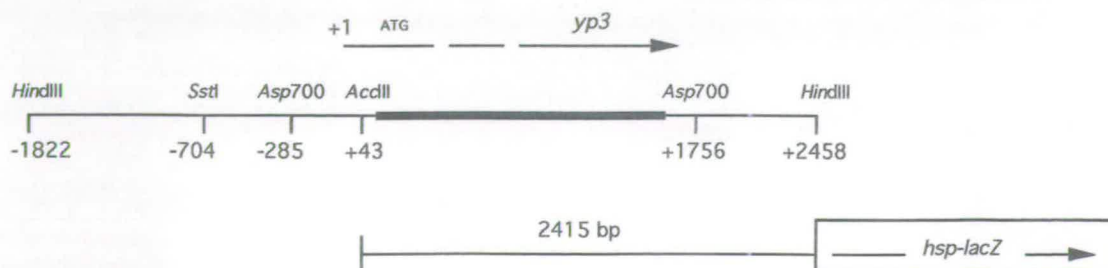


Figure 5.7

**A** Sequence similarity between the 23 bp *Drosophila hsp27* ecdysone response element which has hyphenated dyad symmetry (underlined) and the *yp3* sequence in the 3' flanking DNA

**B** Sequence similarity between the ultraspiracle binding site located upstream of the *Drosophila sl5* chorion gene and the sequence located between the two introns of *yp3*

**C** Sequence similarity between the 23 bp *Drosophila hsp27* ecdysone response element and the sequence located in the *yp3* promoter region

All *yp3* nucleotide positions are given relative to the *yp3* transcription start site and nucleotide identities are indicated by vertical lines



(Martinez et al., 1991). The sequence similarity in the *yp3* coding region was between nucleotides +486 to +495 and is given in Figure 5.7B. The possible 20-hydroxyecdysone response element located in OE3 spanned nucleotides -244 to -266 and also had sequence similarity to the 23 bp sequence found in the *Drosophila hsp27* promoter. This is illustrated in Figure 5.7C. Although these putative sites for 20-hydroxyecdysone response elements associated with *yp3* are supported by experimental evidence, only DNaseI footprinting assays will confirm if they are indeed *yp3* binding sites for the 20-hydroxyecdysone receptor. Until this has been demonstrated, it should also be considered that the response to 20-hydroxyecdysone could be mediated via some other intermediate *trans*-acting factor.

## 5.4 Discussion

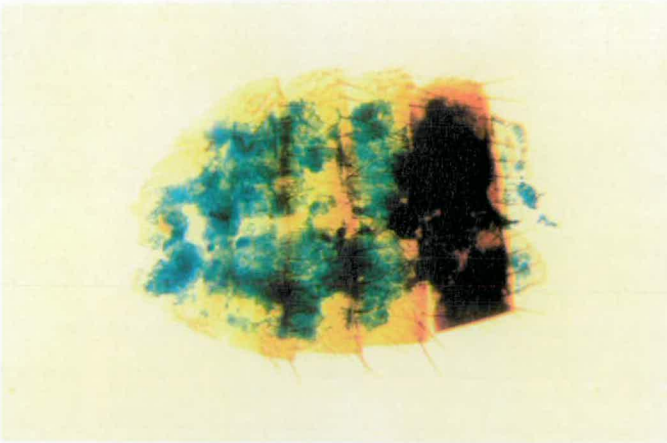
Using DNA regions from the *yp3* 4.3 kb genomic region cloned upstream of the *lacZ* reporter gene, it has been shown that there are ecdysone response elements associated with *yp3* that are involved in the induction of gene expression in response to 20-hydroxyecdysone. It has also been shown that most of the constructs tested have different responses to 20-hydroxyecdysone in the abdominal and thoracic fat body, the two regions where the fat body is located. Based on the results from  $\beta$ -galactosidase assays, a similar response to 20-hydroxyecdysone was only detected with constructs C (the entire upstream region from nts -1822 to +43) and D (the shorter upstream region from nts -704 to +43) in the abdomen and thorax. The most noticeable difference was seen on comparison of the response to 20-hydroxyecdysone induced by constructs D (that contained both the *yp3*FBE and OE3, from nts -704 to +43) and F (that was the FBE on its own, from nts -704 to -285). In the abdominal fat body, levels of reporter gene activity were greater in males transformed with the *yp3*FBE and OE3 region together than in males transformed with the *yp3*FBE alone. However, the reverse was true in the thoracic fat body where higher levels of  $\beta$ -galactosidase activity were observed in males transformed with the *yp3*FBE than in males transformed with the *yp3*FBE and OE3 regions together. Photographs demonstrating this effect are given in Figure 5.8. This could be explained by the presence (in construct D), or absence (in

### Figure 5.8

**A and B**      Photographs representing staining for  $\beta$ -galactosidase activity in the abdominal and thoracic fat body of males transformed with construct D, illustrated below the photographs. This consists of the *yp3FBE* and OE3. Staining is more intense in the abdominal fat body than the thoracic fat body

**C and D**      Photographs representing staining for  $\beta$ -galactosidase activity in the abdominal and thoracic fat body of males transformed with construct F, illustrated below the photographs. This consists of the *yp3FBE* only. Staining is more intense in the thoracic fat body than the abdominal fat body

A

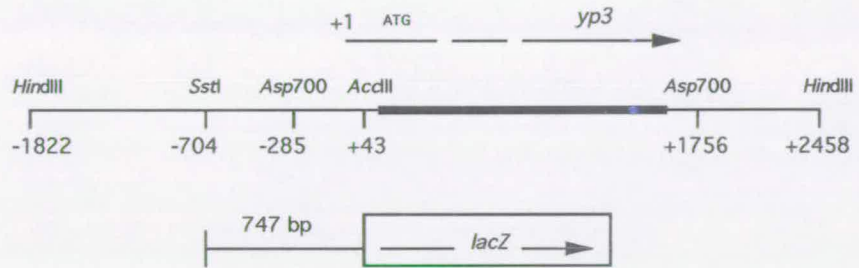


MALE ABDOMEN + 20-HYDROXYECDYSONE

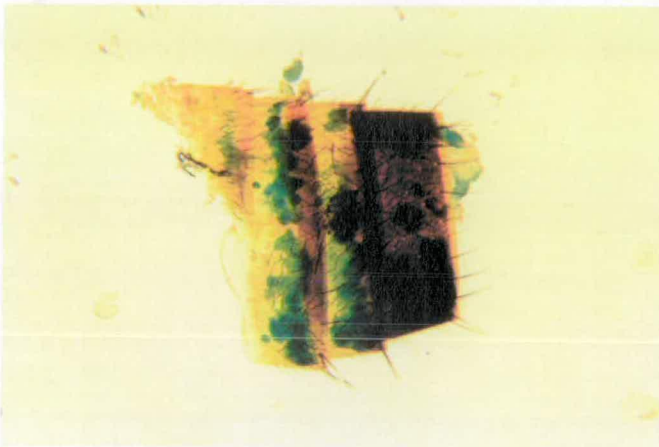
B



MALE THORAX + 20-HYDROXYECDYSONE



C

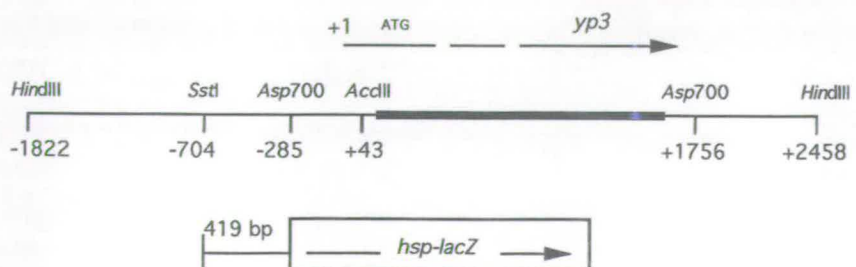


MALE ABDOMEN + 20-HYDROXYECDYSONE

D



MALE THORAX + 20-HYDROXYECDYSONE



construct F) of the *yp3* promoter region (or OE3) as this appears to have different effects in response to 20-hydroxyecdysone in the abdominal and thoracic fat body tissue. In the abdominal fat body the presence of OE3 can enhance levels of reporter gene induction, if the upstream sequences between nt -1822 to -704 are absent. In the thoracic fat body, however, the presence of OE3 can repress levels of reporter gene induction regardless of the presence or absence of upstream sequences spanning the region studied. Both constructs D (*yp3FBE* + OE3) and F (*yp3FBE*) lack the upstream sequence spanning nts -1822 to -704 that appear to suppress a response to 20-hydroxyecdysone in the abdominal fat body. The response in this tissue by construct D, composed of both the *yp3FBE* and OE3 is therefore enhanced. In comparison, the response by construct F, or the *yp3FBE* is reduced due to the absence of OE3. In the thoracic fat body, the presence of OE3 appears to have an opposite, negative, influence. The response to 20-hydroxyecdysone by the construct containing both the *yp3FBE* and OE3 (construct D) is therefore suppressed because of the presence of OE3, and in comparison, the response by the *yp3FBE* (construct F) is much greater due to the absence of OE3. The promoter region of *yp3* spanning nts -285 to +43, also identified as OE3, therefore appears to be responsible for the different responses to 20-hydroxyecdysone observed in the abdominal and thoracic fat body tissue. In the abdominal fat body, the response to 20-hydroxyecdysone is enhanced by the presence of OE3, whereas in the thoracic fat body the hormonal response is suppressed by the presence of OE3.

It therefore appears that there may be differences in the regulation of fat body *yp* gene expression in the abdominal and thoracic fat body regions and it is possible that the different effects of OE3, and the subsequent responses seen in the two fat body locations with most of the constructs tested, could be mediated by different *trans*-acting tissue-specific factors located in the abdomen and thorax, or the same factors present in different concentrations. There is some evidence to suggest that different *trans*-acting factors could be involved in controlling *yp* expression in different subregions of the adult fat body. When a search was carried out for genes involved in the transactivation of *yp* gene expression by screening for mutants that failed to drive expression from a *yp-Adh* gene fusion a number of flies were isolated that had abnormal

fat body development, several of which showed differences in the abdominal and thoracic fat body development (A. Shirras and M. Bownes, personal communication).

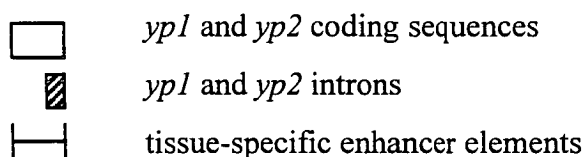


#### 5.4.1 Re-investigation Into the Effects of 20-hydroxyecdysone on *yp1* and *yp2* Expression

The analysis of *yp3* demonstrated that the influence of 20-hydroxyecdysone on reporter constructs resulted from the combined effects of regions which could induce or suppress a response to the hormone. With the investigation of shorter *yp3* constructs, regions which had an enhancing or suppressing influence were separated and their effects more easily monitored. In the initial analysis of *yp1* and *yp2* (Bownes and Shirras, unpublished results; Shirras and Bownes, 1987) the constructs tested contained most of the upstream sequences. Further analysis of the *yp1* and *yp2* genomic region was therefore undertaken to determine if 20-hydroxyecdysone responsive elements could be identified with shorter regions cloned upstream of reporter constructs (Bownes et al., submitted). These experiments were performed by D.Mauchline and are included for comparison to *yp3*. The constructs tested are illustrated in Figure 5.9. Recent data from this analysis showed that males transformed with shorter regions from the *yp1* and *yp2* intergenic region fused to an *Adh* reporter gene (Figure 5.9, constructs A-F) also responded to 20-hydroxyecdysone. Construct A was a 929 bp *yp1* fragment spanning nucleotides -886 to +43 previously shown to have no response to 20-hydroxyecdysone (Shirras and Bownes, 1987). Deletions made in this region, however, did lead to 20-hydroxyecdysone responsiveness. The deletions were mainly around the 125 bp fat body enhancer (FBE) located between nucleotides -196 and -321 upstream of the *yp1* transcription start site. The FBE itself did not confer 20-hydroxyecdysone responsiveness in transformed males (construct I) but when it was deleted from the 929 bp *yp1* fragment, transformed males responded to treatment with 20-hydroxyecdysone (constructs C and D). Construct B partially deleted the FBE and occasionally had a weak response to 20-hydroxyecdysone, suggesting that a response element may have been partially lost. However when the FBE was present in construct F spanning nucleotides -321 to +38 and the other flanking sequences were deleted (-886 to -321), responsiveness to 20-hydroxyecdysone was maintained. Other constructs using  $\beta$ -galactosidase as the reporter gene were also investigated (constructs G, H and I) but did not respond to 20-hydroxyecdysone. Construct J was the final construct tested

Figure 5.9

Constructs used in the detailed analysis to identify 20-hydroxyecdysone (20-OH) response elements associated with *yp1* and *yp2*. The genomic arrangement of *yp1* and *yp2* is shown at the top of the diagram with the restriction enzyme sites used to make some of the constructs shown below

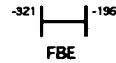
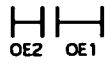


Transcripts are represented as arrows with the arrowhead representing the direction of transcription. The alcohol dehydrogenase (*Adh*) and  $\beta$ -galactosidase (*hsp70-lacZ*) reporter genes are shown with the appropriate constructs. All nucleotide positions are given in relation to the *yp1* transcription start site (+1)

+ indicates expression

- indicates no expression

+/- indicates expression sometimes



FEMALE  
FAT BODY  
STAINING

20-OH  
INDUCTION  
IN MALES

A		+	-
B		+	+/-
C		+	+
D		+	+
E		+	-
F		+	+
G		+	-
H		+	-
I		+	-
J		+	+

in which a response to 20-hydroxyecdysone was observed. This contained the coding region of *yp2* with 343 bp of 5' flanking sequences, the induction of which was monitored in a *yp2*<sup>-</sup> background.

From these analyses, it was concluded that at least two 20-hydroxyecdysone response elements were associated with the *yp1* and *yp2* genes, one of which was located in the 5' sequences flanking *yp1* and one in either coding region or sequences upstream of *yp2*. It also appeared that multiple sites could function to repress reporter gene transcription in response to stimulation by 20-hydroxyecdysone, and deletion of some of these sites allowed transcription to proceed. The results of this analysis also indicated that 20-hydroxyecdysone response elements were separate from, but interacted with, nearby repressor binding sites. It is possible that the multiple DSX binding sites present in the FBE were responsible for suppressing a response to 20-hydroxyecdysone since the FBE did not respond to 20-hydroxyecdysone itself, and prevented a response by the 929 bp *yp1* fragment when it was included in this construct. However there is no direct evidence to prove this and other factors could be responsible for this observed effect, such as *trans*-acting tissue specific factors acting as repressors of gene activation. It has also been reported that ecdysterone regulatory elements can function as both transcriptional activators and repressors, depending on the number of elements that are present (Dobens et al., 1991). This analysis was carried out using the 23 bp ecdysone response element found in the *Drosophila hsp27* promoter (Riddiough and Pelham, 1987), which functioned as a gene activator in response to 20-hydroxyecdysone when one or two of the 23 bp sequences were tested. When these sequences were tested in tandem repeats (10-mers or 40-mers) the response to 20-hydroxyecdysone was reduced, demonstrating that these sequences could also act as repressors of reporter gene induction. It is not known whether these observations reflect competition for transcription factors involved in mediating a response to 20-hydroxyecdysone, although it has been reported that steroid hormone receptors can compete for factors that mediate their enhancer function (Meyer et al., 1989). In analogy to this, it is possible that 20-hydroxyecdysone response elements were simply separated in the analysis of the *yps*, thereby allowing their activating potential to be monitored.

#### 5.4.2 The Effects of Juvenile Hormone on *yp* Transcription

Most of the *yp* constructs used to identify the presence of *cis*-acting 20-hydroxyecdysone response elements were also assessed for the presence of juvenile hormone (JH) response elements (D.Mauchline and M.Bownes, personal communication). Since JH treatment of males does not induce expression of the *yps*, the effects of this hormone must be monitored in females. Females were starved to reduce background levels of *yp* transcription with reporter constructs such that any effects of JH could be monitored. An increase in reporter gene expression was not observed with any of the constructs tested. A response to JH only occurred with the native *yp* genes, which was monitored for *yp2* and *yp3* using transgenes in mutants that lacked YP2 and YP3 in the haemolymph and eggs. Levels of transgenic *yp2* and *yp3* expression were reduced in starved females, and increased in JH-treated flies. This effect was also seen with the native *yp* genes. It therefore seems that JH acts upon *yp* transcripts, either in the coding sequences or introns, perhaps to influence transcript stability. This is consistent with recent investigations into the role of JH in *Manduca sexta*, the tobacco hornworm. A JH binding protein was recently identified, thought to be the JH receptor. Although it was localised to the nucleus, sequence analysis revealed that it had no known DNA-binding motifs. However, it did have sequence properties similar to an RNA helicase, and was also present in structures associated with the nucleolus. It has therefore been postulated to influence RNA metabolism by affecting post-transcriptional RNA processing or RNA translation (Palli et al., 1994). It is therefore possible that an equivalent to the *Manduca* JH receptor is present in *Drosophila*, and may influence *yp* expression by affecting transcript stability.

## **CHAPTER 6**

### **Final Discussion**

## 6.1 Enhancer Regions Required for Tissue-Specific *yp3* Expression

The *yp3* gene of *Drosophila melanogaster* was investigated with the aim of localising *cis*-acting regulatory elements involved in controlling the tissue-, sex- and temporal pattern of expression. Various upstream, coding and downstream regions, and combinations of these, were tested for their effects on *lacZ* reporter gene expression *in vivo* by P-element mediated germ line transformation and assaying for  $\beta$ -galactosidase activity by histological staining. Using this technique, two separate and independently acting *cis*-regulatory regions were identified that were sufficient for the tissue-specific expression of *yp3* in the ovary and fat body. The fat body enhancer (*yp3*FBE) was localised to a 419 bp fragment spanning nucleotides -704 to -285 and was also sufficient to maintain sex-limited expression, as no  $\beta$ -galactosidase activity was detected in the fat body of adult males transformed with this construct. The ovarian enhancer (termed OE3) was contained within a neighbouring 328 bp promoter fragment spanning nucleotides -285 to +43 (Ronaldson and Bownes, in press). No other sequences were identified which were autonomously involved in determining the *yp3* expression pattern, although further analysis would be required to determine whether the identified enhancer regions interact with sequences contained within the *yp3* coding or 3' regions. It also remains to be established whether OE3 and the *yp3*FBE themselves contain a number of elements that interact to direct the sex-, tissue- and developmental-specificity of *yp* expression, which will require the further localisation and precise identification of internal enhancer sequences.

It appears that the transcription of many genes with a tissue- and/or developmentally- regulated pattern of expression is mediated by competition between *cis*-acting activator and repressor sites (Müller and Bienz, 1991; Busturia and Bienz, 1993), and this appears to be part of the mechanism involved in controlling the overlapping expression of the yolk protein (stage 8-10), vitelline membrane (stage 8-10) and chorion genes (stage 10b-14) in the ovarian follicle cells (Mariani et al., 1988; Tolias and Kafatos, 1990; Bienz-Tadmor et al., 1992; Jin and Petri, 1993).

## 6.2 ***Trans-acting Factors Mediating yp3 Fat Body Transcription***

Computer searches for the binding motifs of *trans*-acting factors thought to be involved in gene regulation can be used to indicate the location of important regulatory sequences. Putative binding sites for the fat body transcription factors BBF-2 (Abel et al., 1992), AEF-1 and C/EBP (Falb and Maniatis, 1992b), that bind to the *yp1/yp2* FBE *in vitro*, have been identified in the *yp3*FBE using computer searches with the footprints obtained from *yp1/yp2* FBE binding. Several potential DSX binding sites (Burtis et al., 1991) are also present. The outcome of these computer searches could be helpful in the design of future reporter constructs. Further dissection of the *yp3*FBE followed by P-element mediated germ-line transformation studies would separate any internal enhancer sequences and reveal the significance of the computer searches. These studies could be complemented by *in vitro* gel retardation assays with binding factors and the *yp3*FBE and sub-fragments. Preliminary gel retardation assays with the DNA binding domain of the DSX protein have shown that it specifically binds to the *yp3*FBE *in vitro* (J.Richardson and M.Bownes, personal communication).

## 6.3 ***Trans-acting Factors Mediating Ovarian yp3 Transcription***

Gel retardation assays were performed with a 150 bp sub-fragment (-197 to -47) termed OE3F, derived from OE3 by PCR to remove the TATA and CAAT-box promoter elements. It was demonstrated that ovarian nuclear extracts contained a protein or protein complex that specifically bound to the 150 bp sequence, which could be novel proteins involved in the regulation of ovarian *yp* transcription since computer searches for the binding sites of known transcription factors in the 150 bp sequence revealed no obvious sequence similarities. It would be of great interest to identify the protein(s) involved in OE3F binding, since there have been no reports on the isolation of factors involved in the regulation of ovarian *yp* transcription. This could be achieved using OE3F to screen a *Drosophila* ovarian cDNA expression library to obtain a full-length cDNA clone encoding the DNA binding protein (Shea et al., 1990; Singh, 1993). However the success of this approach would depend on such a protein having the ability



to recognise its associated binding site in the absence of other factors that may be involved in the formation of a transcription complex *in vivo*. It would therefore be necessary to adopt other approaches, one of which could be to purify proteins directly from ovarian nuclear extracts by means of Magnetic DNA Affinity Purification (Gabrielsen and Huet, 1993). With this system, the DNA binding fragment is biotinylated at one end such that it forms a strong bond with streptavidin coated magnetic dynabeads. On incubation with the ovarian nuclear extract, any proteins of interest will bind their target sequence and the entire complex can be magnetically separated from other proteins in the reaction. This allows the isolation of factors that may not be involved in DNA recognition but are required for protein-protein interactions in the co-operative formation of a DNA binding transcription complex. Once separated, proteins are eluted from the DNA binding fragment and separated by SDS polyacrylamide gel electrophoresis. Discrete bands can then be used directly for peptide sequencing and used in preliminary database searches for homology to known proteins. If sequences are novel they can provide a basis for the design of oligonucleotides with which to screen a *Drosophila* ovarian cDNA library and isolate genes encoding proteins governing ovarian *yolk protein* transcription. In the long term, the availability of genes encoding such transcription factors would facilitate site-directed mutagenesis experiments to identify important structural or functional domains, the ultimate aim of which would be to understand their mode of action. Assuming a full length cDNA clone can be identified that encodes a protein capable of independently binding OE3F, it could be expressed and purified for *in vitro* footprinting to precisely locate the binding sequence. Footprinting could also be attempted using ovarian nuclear extracts since the formation of a protein complex may be a prerequisite for recognition and association with the DNA binding site (Shea et al., 1990; Frank et al., 1992).

#### **6.4 The Hormonal Control of *yp* Transcription**

A topical application of 20-hydroxyecdysone by injection into male abdomens leads to transient expression of the *yps* in the fat body tissue (Postlethwait et al., 1980; Bownes, 1982). The effect of 20-hydroxyecdysone on *yp3* gene expression

was therefore investigated using reporter constructs containing various flanking sequences, coding sequences and combinations of these fused upstream of the *lacZ* reporter gene. *Cis*-acting regions required for the induction of *yp3* transcription in response to 20-hydroxyecdysone were mapped to sites located upstream, downstream and within the coding sequence or introns of *yp3*. It appeared that a site suppressing a response to 20-hydroxyecdysone was present in the 5' flanking region, as a shorter upstream construct consistently had a greater response to 20-hydroxyecdysone than a construct containing all of the upstream sequences. It was also observed that the downstream sequences and the coding region, or introns, interacted to produce a response to 20-hydroxyecdysone, or that a 20-hydroxyecdysone response element was located at the border separating these regions.

The influence of 20-hydroxyecdysone on the *yp1* and *yp2* genes was also investigated (D. Mauchline and M. Bownes) the results of which were included for comparison to the *yp3* studies. As with *yp3*, 20-hydroxyecdysone appeared to directly influence *yp* expression as *cis*-acting sites were localised to 5' sequences flanking *yp1* and to either the *yp2* coding or upstream sequences (Bownes et al., submitted). In summary, from analysis of the effects of 20-hydroxyecdysone on various *yp1*, *yp2* and *yp3* reporter constructs, it appeared that complex interactions were occurring between 20-hydroxyecdysone response elements and repressor sites, to determine the response to this hormone. Although it appeared that 20-hydroxyecdysone could be affecting *yp* expression directly, by interacting with ecdysone receptor molecules in association with the ultraspiracle (USP) protein, it was also possible that another intermediate gene or genes were involved in this effect. Evidence for this came from the investigations with *yp3*, where there was a difference in the response to 20-hydroxyecdysone in the thoracic and abdominal fat body between different constructs in transformed males, suggesting that different *trans*-acting factors could be mediating the response in sub-regions of the adult fat body. If this is the case, then the *trans*-acting factor(s) responsible will need to be identified to continue investigations into the hormonal regulation of *yp* gene expression. Since USP (Shea et al., 1990) and the 20-hydroxyecdysone receptor have been cloned (Koelle et al., 1991; Talbot et al., 1993) and antibodies are available to the different ecdysone receptor isoforms, DNA-binding studies are currently underway, to

determine if the effects of 20-hydroxyecdysone on *yp* transcription are direct. If this is the case, it will be possible to investigate the interactions that may be occurring between the 20-hydroxyecdysone receptor complex, the putative fat body *trans*-acting factors (Abel et al., 1992; Falb and Maniatis, 1992) and the DSX<sup>M</sup> and DSX<sup>F</sup> proteins (Burtis et al., 1991).

In comparison, no sites were identified to indicate that there was a direct interaction between JH and the *yp* genes and it is therefore thought that JH may act upon *yp* transcripts rather than DNA sequences to influence levels of expression. A JH receptor that may mediate this effect remains to be identified in *Drosophila melanogaster*.

## 6.5 Concluding Remarks

The *yolk protein* genes of *Drosophila melanogaster* have provided an ideal system to identify factors involved in the transcriptional control mediating their highly specific pattern of sex-limited, tissue-specific and developmentally coordinated expression. Enhancer regions have been localised for all three genes that are required for sex-specific expression in the adult female fat body, and transcription in the ovarian follicle cells during limited stages of oogenesis. Regions of sequence similarity exist between the fat body elements and between the identified ovarian enhancers, and are likely to contain putative recognition sites for *trans*-acting factors. Several *trans*-acting factors have been implicated in the control of *yp* transcription in the fat body from *in vitro* footprinting assays to the *yp1/yp2* FBE, and similar studies are underway with *yp3*.

It would be worthwhile pursuing the isolation of ovarian *trans*-acting factors using the identified ovarian enhancer for *yp3* since none have been identified to date, and any ovarian transcription factors identified would be expected to be involved in the transcriptional regulation of all three *yp* genes. Ultimately, if all the *trans*-acting factors involved in the regulation of *yolk protein* gene transcription can be identified and isolated, it will be possible to gain further knowledge concerning the interactions between them, and hence the mechanisms involved in the transcriptional control of this small gene family.

It has also become clear that there is a direct hormonal influence on *yp* transcription, resulting from a combination of positive and negative influences by *cis*-acting sequences associated with the *yps*. The significance of this in adult females *in vivo* is difficult to determine, but a hormonal regulation of gene expression would provide further scope for the regulation or maintenance of a highly specific pattern of gene expression.

**REFERENCES**

- Abel, T., Bhatt, R. and Maniatis, T.** (1992) A *Drosophila* CREB/ATF transcriptional activator binds to both fat body- and liver-specific regulatory elements. *Genes Dev.* 6: 466-480
- Abrahamsen, N., Martinez, A., Kjær, T., Søndergaard, L. and Bownes, M.** (1993) Cis-regulatory sequences leading to female-specific expression of yolk protein genes 1 and 2 in the fat body of *Drosophila melanogaster*. *Mol. Gen. Genet.* 237: 41-48
- Amin, J., Mestril, R., Lawson, R., Klapper, H. and Voellmy, R.** (1985) The heat shock consensus sequence is not sufficient for hsp70 gene expression in *Drosophila melanogaster*. *Mol. Cell. Biol.* 5: 197-203
- Andres, A.J. and Thummel, C.S.** (1992) Hormones, puffs and flies: the molecular control of metamorphosis by ecdysone. *Trends Genet.* 8: 132-138
- Ashburner, M. and Richards, G.** (1976) Sequential gene activation by ecdysone in polytene chromosomes of *Drosophila melanogaster*. III. Consequences of ecdysone withdrawal. *Dev. Biol.* 54: 241-255
- Ashburner, M., Chihara, C., Meltzer, P. and Richards, G.** (1974) Temporal control of puffing activity in polytene chromosomes. *Cold Spring Harbour Symp. Quant. Biol.* 38: 655-662
- Baeuerle, P.A. and Huttner, W.B.** (1985) Tyrosine sulfation of yolk proteins 1, 2 and 3 in *Drosophila melanogaster*. *J. Biol. Chem.* 260: 6434-6439
- Baeuerle, P.A., Lottspeich, F. and Huttner, W.B.** (1988) Purification of yolk protein 2 of *Drosophila melanogaster*. *J. Biol. Chem.* 263: 14925-14929
- Baker, B.S.** (1989) Sex in flies: the splice of life. *Nature* 340: 521-524
- Baker, B.S. and Belote, J.M.** (1983) Sex determination and dosage compensation in *Drosophila melanogaster*. *Ann. Rev. Genet.* 17: 345-393
- Bandziulis, R.J., Swanson, M.S. and Dreyfuss, G.** (1989) RNA-binding proteins as developmental regulators. *Genes Dev.* 3: 431-437
- Baniahmad, A., Muller, M., Steiner, C. and Renkawitz, R.** (1987) Activity of two different silencer elements of the chicken lysozyme gene can be compensated by enhancer elements. *EMBO J.* 6: 2297-2303

- Barnett, T. and Wensink, P.C.** (1981) Transcription and translation of yolk protein mRNA in the fat bodies of *Drosophila*. In: Developmental Biology using purified genes, eds. Brown, D.D and Fox, C.F. UCLA-ICN Symposium on Molecular and Cellular Biology. University of California Press, Berkeley 23: pp 99-106
- Barnett, T., Pachl, C., Gergen, J.P. and Wensink, P.C.** (1980) The isolation and characterisation of *Drosophila* yolk protein genes. *Cell* 21: 729-738
- Bell, L.R., Maine, E.M., Schedl, P. and Cline, T.W.** (1988) *Sex-lethal*, a *Drosophila* sex-determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. *Cell* 55: 1037-1046
- Bell, L.R., Horabin, J.I., Schedl, P. and Cline, T.W.** (1991) Positive autoregulation of *Sex-lethal* by alternative splicing maintains the female determined state in *Drosophila*. *Cell* 65: 229-239
- Belote, J.M.** (1989) The control of sexual development in *Drosophila melanogaster*: genetic and molecular analysis of a genetic regulatory hierarchy - a minireview. *Gene* 82: 161-167
- Belote, J.M.** (1992) Sex determination in *Drosophila melanogaster*: from the X:A ratio to *doublesex*. *Semin. Dev. Biol.* 3: 319-330
- Belote, J.M. and Baker, B.S.** (1982) Sex determination in *Drosophila melanogaster*: analysis of *transformer-2*, a sex-transforming locus. *Proc. Natl. Acad. Sci. USA* 79: 1568-1572
- Belote, J.M. and Baker, B.S.** (1983) The dual functions of a sex-determination gene in *Drosophila melanogaster*. *Dev. Biol.* 95: 512-517
- Belote, J.M., Handler, A.M., Wolfner, M.F., Livak, K.J. and Baker, B.S.** (1985) Sex-specific regulation of yolk protein gene expression in *Drosophila*. *Cell* 40: 339-348
- Bienz-Tadmor, B., Tolia, P., Stebbins-Boaz, B., Mariani, B.D., Gerbi, S.A. and Kafatos, F.C.** (1992) Chorion gene *cis*-regulatory DNA restricts tissue specificity of reporter gene expression in transformed *Drosophila*. *Chromosoma* 101: 538-548

- Boddine, D.M. and Ley, T.J.** (1987) An enhancer element lies 3' to the human A<sub>γ</sub> globin gene. *EMBO J.* 6: 2997-3004
- Bownes, M.** (1975) A photographic study of development in the living embryo of *Drosophila melanogaster*. *J. Embryol. Exp. Morph.* 33: 789-801
- Bownes, M.** (1982) Hormonal and genetic regulation of vitellogenesis in *Drosophila*. *The Quarterly Review of Biology* 57: 247-274
- Bownes, M.** (1986) Expression of the genes coding for vitellogenins (Yolk Proteins). *Ann. Rev. Entomol.* 31: 507-531
- Bownes, M.** (1989) The roles of juvenile hormone, ecdysone and the ovary in the control of *Drosophila* vitellogenesis. 35: 409-413
- Bownes, M.** (1990) The nature and origin of ecdysteroids during *Drosophila* metamorphosis. In: *Molting and Metamorphosis*. Ed. E. Ohnishi and H. Ishizahi. Japan Scientific Societies Press and Springer-Verlag, Tokyo. pp 107-120
- Bownes, M.** (1994) Interactions between germ cells and somatic cells in *Drosophila melanogaster*. *seminars in Dev. Biol.* 5: 31-42
- Bownes, M. and Blair, M.** (1986) The effect of a sugar diet and hormones on the expression of the *Drosophila* yolk protein genes. *J. Insect Physiol.* 32: 493-501
- Bownes, M. and Dale, L.** (1982) Gametogenesis. In: *Handbook of Drosophila development*. Ed. R. Ransom, Elsevier Biochemical Press, Oxford. pp 31-66
- Bownes, M. and Hames, B.D.** (1977) Accumulation and degradation of three major yolk proteins in *Drosophila melanogaster*. *J. Exp. Zool.* 200: 149-156
- Bownes, M. and Hames, B.D.** (1978) Analysis of the yolk proteins in *Drosophila melanogaster*. *FEBS Lett.* 96: 327-330
- Bownes, M. and Nöthiger, R.** (1981) Sex determining genes and vitellogenin synthesis in *Drosophila melanogaster*. *Mol. Gen. Genet.* 182: 222-228
- Bownes, M. and Rembold, H.** (1987) The titre of juvenile hormone during the pupal and adult stages of the life cycle of *Drosophila melanogaster*. *FEBS* 164: 709-712



- Bownes, M., Blair, M., Kozma, R. and Dempster, M.** (1983) 20-hydroxyecdysone stimulates tissue-specific yolk protein gene transcription in both male and female *Drosophila*. *J. Embryol. Exp. Morph.* 78: 249-268
- Bownes, M., Scott, A. and Blair, M.** (1987) The use of an inhibitor of protein synthesis to investigate the roles of ecdysteroids and sex-determination genes on the expression of the genes encoding the *Drosophila* yolk proteins. *Development* 101: 931-941
- Bownes, M., Scott, A. and Shirras, A.** (1988a) Dietary components modulate yolk protein gene transcription in *Drosophila melanogaster*. *Development* 103: 119-128
- Bownes, M., Shirras, A., Blair, M., Collins, J. and Coulson, A.** (1988b) Evidence that insect embryogenesis is regulated by ecdysteroids released from yolk proteins. *Proc. Natl. Acad. Sci. USA.* 85: 1554-1557
- Bownes, M., Ronaldson, E., Mauchline, D. and Martinez, A.** (1993) Regulation of vitellogenesis in *Drosophila*. *Int. J. Insect Morphol. Embryol.* 22: 349-367
- Bownes, M., Ronaldson, E. and Mauchline, D.** (1995) 20-hydroxyecdysone, but not juvenile hormone, regulation of *yolk protein* gene expression can be mapped to flanking cis-acting DNA sequences. Submitted for publication in *Development*.
- Bradford, M.M.** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding. *Anal. Biochem.* 72: 248-254
- Brennan, M.D., Warren, T.G. and Mahowald, A.P.** (1980) Signal peptides and signal peptidases in *Drosophila melanogaster*. *J. Cell Biol.* 87: 516-520
- Brennan, M.D. and Mahowald, T.G.** (1982) Phosphorylation of the vitellogenin polypeptides of *Drosophila melanogaster*. *Insect Biochem.* 12: 669-673
- Brennan, M.D., Weiner, A.J., Goralski, T.J. and Mahowald, A.P.** (1982) The follicle cells are a major site of vitellogenin synthesis in *Drosophila melanogaster*. *Dev. Biol.* 89: 225-236
- Briggs, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R.** (1986) Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* 234: 47

- Burke, T., Waring, G.L., Popodi, E. and Minoo, P.** (1987) Characterization and sequence of follicle cell genes selectively expressed during vitelline membrane formation in *Drosophila melanogaster*. *Dev. Biol.* 124: 441-450
- Burtis, K.C. and Baker, B.S.** (1989) *Drosophila doublesex* gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* 56: 997-1010
- Burtis, K.C., Thummel, C.S., Jones, C.W., Karim, F.D and Hogness, D.S.** (1990) The *Drosophila* 74EF early puff contains E74, a complex ecdysone -inducible gene that encodes two *ets*-related proteins. *Cell* 61: 85-99
- Burtis, K.C., Coschigano, K.T., Baker, B.S. and Wensink, P.C.** (1991) *Drosophila doublesex* protein binds to a sex-specific yolk protein gene enhancer. *EMBO J.* 10: 2577-2582
- Busturia, A. and Bienz, M.** (1993) Silencers in *Abdominal-B*, a homeotic *Drosophila* gene. *EMBO J.* 12: 1415-1425
- Butterworth, F.M., Bownes, M. and Burde, V.S.** (1991) Genetically modified yolk proteins precipitate in the adult *Drosophila* fat body. *J. Cell Biol.* 112: 727-737
- Cabrera, C.V. and Alonzo, M.C.** (1991) Transcriptional activation by heterodimers of the *achaetae-scute* and *daughterless* gene products of *Drosophila*. *EMBO J.* 10: 2965-2973
- Caudy, M., Vaissin, H., Brand, M., Tuma, R., Jan, L.Y. and Jan, Y.N.** (1988) *daughterless*, a *Drosophila* gene essential for both neurogenesis and sex-determination, has sequence similarities to *myc* and the *achaete-scute* complex. *Cell* 55: 1061-1067
- Cherbas, L., Lee, K. and Cherbas, P.** (1991) Identification of ecdysone response elements by analysis of the *Drosophila Eip28/29* gene. *Genes Dev.* 5: 120-131
- Choi, E-R.B. and Engel, J.D.** (1988) Developmental regulation of  $\beta$ -globin gene switching. *Cell* 55: 17-26
- Christianson, A.M.K., King, D.L., Hatzivassiliou, E., Casas, J.E., Hallenbeck, P.L., Nikodem, V.M., Mitsialis, S.A. and Kafatos, F.C.** (1992) DNA binding and heteromerization of the *Drosophila* transcription factor chorion factor 1/ultraspiracle. *Proc. Natl. Acad. Sci. USA* 89: 11503-11507

- Christianson, A.M.K. and Kafatos, F.C.** (1993) Binding affinity of the *Drosophila melanogaster* CF1/ USP protein to the chorion *s15* promoter. Biochem. Biophys. Res. Comm. 193: 1318-1323
- Cline, T.W.** (1978) Two closely linked mutations in *Drosophila melanogaster* that are lethal to opposite sexes and interact with *daughterless*. Genetics 90: 683-698
- Cline, T.W.** (1983) The interaction between *daughterless* and *Sex-lethal* in triploids: A lethal sex-transforming maternal effect linking sex-determination and dosage compensation in *Drosophila melanogaster*. Dev. Biol. 95: 260-274
- Cline, T.W.** (1984) Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state. Genetics 107: 231-277
- Cline, T.W.** (1993) The *Drosophila* sex determination signal: How do flies count to two? TIGS 9: 385-390
- Coschigano, K.T. and Wensink, P.C.** (1993) Sex-specific transcriptional regulation by the male and female doublesex proteins of *Drosophila*. Genes Dev. 7: 42-54
- Cronmiller, C., Schedl, P. and Cline, T.W.** (1988) Molecular characterisation of *daughterless*, a *Drosophila* sex-determination gene with multiple roles in development. Genes Dev. 2: 1666-1676
- Devereux, J., Haeblerli, P. and Smithies, O.** (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12: 387-395
- DiBello, P.R., Withers, D.A., Bayer, C.A., Fristrom, J.W. and Guild, G.M.** (1991) The *Drosophila Broad-Complex* encodes a family of related, zinc finger-containing proteins. Genetics 129: 385-397
- Dobens, L., Rudolph, K. and Berger, E.M.** (1991) Ecdysone regulatory elements function as both transcriptional activators and repressors. Mol. Cell. Biol. 11: 1846-1853
- Dudler, R. and Travers, A.A.** (1984) Upstream elements necessary for optimal function of the hsp70 promoter in transformed flies. Cell 38: 391-398
- Erdman, S.E. and Burtis, K.C.** (1993) The *Drosophila* doublesex proteins share a novel zinc finger related DNA binding domain. EMBO J. 12: 527-535
- Erickson, J.W. and Cline, T. W.** (1993) A bZIP protein, sisterless-a, collaborates with bHLH transcription factors early in *Drosophila* development to determine sex. Genes Dev. 7: 1688-1702

- Evans, R.M.** (1988) The steroid and thyroid hormone receptor superfamily. *Science* 240: 889-895
- Feinberg, A.P. and Vogelstein, B.** (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132: 6-13
- Feinberg, A.P. and Vogelstein, B.** (1984) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Addendum. *Anal. Biochem.* 137: 226-267
- Falb, D. and Maniatis, T.** (1992a) *Drosophila* transcriptional repressor protein that binds specifically to negative control elements in fat body enhancers. *Mol. Cell. Biol.* 12: 4093-4103
- Falb, D. and Maniatis, T.** (1992b) A conserved regulatory unit implicated in tissue-specific gene expression in *Drosophila* and man. *Genes Dev.* 6: 454-465
- Fargnoli, J. and Waring, G.L.** (1982) Identification of vitelline membrane proteins in *Drosophila melanogaster*. *Dev. Biol.* 92: 306-314
- Fischer, J.A. and Maniatis, T.** (1988) *Drosophila Adh*: a promoter element expands the tissue specificity of an enhancer. *Cell* 53: 451-461
- Flickinger, T.W. and Salz, H.K.** (1994) The *Drosophila* sex determination gene *snf* encodes a nuclear protein with sequence and functional similarity to the mammalian U1A snRNP protein. *Genes Dev.* 8: 914-925
- Forman, B.M. and Samuels, H.J.** (1990) Interactions among a sub-family of nuclear hormone receptors: The regulatory zipper model. *Mol. Endocrinol.* 4: 1293-1301
- Fourney, R., Pratt, G., Harnish, D., Wyatt, G. and White, B.** (1982) Structure and synthesis of vitellogenin and vitellin from *Calliphora erythrocephala*. *Insect Biochem.* 12: 311-321
- Frank, L.H., Cheung, H-K. and Cohen, R.S.** (1992) Identification and characterization of *Drosophila* female germ line transcriptional control elements. *Development* 114: 481-491
- Friederich, E., Baeuerle, P.A., Garoff, H., Hovemann, B. and Huttner, W.B.** (1988) Expression, tyrosine sulfation and secretion of yolk protein 2 of *Drosophila melanogaster* in mouse fibroblasts. *J. Biol. Chem.* 263: 14930-14938

- Gabrielsen, O.S. and Huet, J.** (1993) *Meths. Enzymol.* 218: 508-525
- Garabedian, M.J., Hung, M-C. and Wensink, P.C.** (1985) Independent control elements that determine yolk protein gene expression in alternative *Drosophila* tissues. *Proc. Natl. Acad. Sci. USA* 82: 1396-1400
- Garabedian, M.J., Shepherd, B.M. and Wensink, P.C.** (1986) A tissue specific transcription enhancer from the *Drosophila* yolk protein 1 gene. *Cell* 45: 859-867
- Garabedian, M.J., Shirras, A.D., Bownes, M. and Wensink, P.C.** (1987) The nucleotide sequence of the gene coding for *Drosophila melanogaster* yolk protein 3. *Gene* 55: 1-8
- Gergen, J.P.** (1987) Dosage compensation in *Drosophila*: evidence that *daughterless* and *Sex-lethal* control X chromosome activity at the blastoderm stage of embryogenesis. *Genetics* 117: 477-485
- Geyer, P.K. and Corces, V.G.** (1992) DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes Dev.* 6: 1865-1873
- Gigliotti, S., Graziani, F., De Ponti, L., Rafti, F., Manzi, A., Lavorgna, G., Gargiulo, G. and Malva, C.** (1989) Sex-, tissue- and stage-specific expression of a vitelline membrane protein gene from region 32 of the second chromosome of *Drosophila melanogaster*. *Dev. Genet.* 10: 33-41
- Giorgi, F., Lucchesi, P., Morelli, A. and Bownes, M.** (1993) Ultrastructural analysis of *Drosophila* ovarian follicles differing in yolk polypeptide (yps) composition. *Development* 117: 319-328
- Gochoco, C.H., Kunkel, J.G. and Nordin, J.H.** (1988) Experimental modification of an insect vitellin affects its structure and its uptake by oocytes. *Arch. Insect Biochem. Physiol.* 2: 179-199
- Goodbourn, S., Burstein, H. and Maniatis, T.** (1986) The human  $\beta$ -interferon gene enhancer is under negative control. *Cell* 45: 601-610
- Goralski, T.J., Edström, J.E. and Baker, B.S.** (1989) The sex-determination locus *transformer-2* of *Drosophila* encodes a polypeptide with similarity to RNA binding proteins. *Cell* 56: 1011-1018

- Gough, J. and Murray, N.** (1983) Sequence diversity among related genes for recognition of specific targets in DNA molecules. *J. Mol. Biol.* 166: 1-19
- Griffin-Shea, R., Thireos, G. and Kafatos, F.C.** (1982) Organization of a cluster of four chorion genes in *Drosophila* and its relationship to developmental expression and amplification. *Dev. Biol.* 91: 325-336
- Hedley, M.L. and Maniatis, T.** (1991) Sex-specific splicing and polyadenylation of *dsx* pre-mRNA requires a sequence that binds specifically to *tra-2* protein. *Cell* 65: 579-586
- Higgins, M.J., Walker, V.K., Holden, J.J. and White, B.N.** (1984) Isolation of two *Drosophila melanogaster* genes abundantly expressed in the ovary during vitelline membrane synthesis. *Dev. Biol.* 105: 155-165
- Hilfiker, A. and Nöthiger, R.** (1991) The temperature-sensitive mutation *vir* (*virilizer*) identifies a new gene involved in sex determination of *Drosophila*. *Roux's Arch. Dev. Biol.* 200: 240-248
- Hoshijima, K., Inoue, K., Higuchi, I., Sakamoto, H. and Shimura, Y.** (1991) Control of *doublesex* alternative splicing by *transformer* and *transformer-2* in *Drosophila*. *Science* 252: 833-836
- Hovemann, B., Galler, R., Walldorf, U., Küpper, H. and Bautz, E.K.F.** (1981) Vitellogenin in *Drosophila melanogaster*: sequence of the yolk protein 1 gene and its flanking regions. *Nuc. Acids Res.* 9: 4721-4734
- Hu, L. and Gudas, L.J.** (1994) Activation of keratin 19 gene expression by a 3' enhancer containing an AP1 site. *J. Biol. Chem.* 269: 183-191
- Huff, C.A., Yuspa, S.H. and Rosenthal, D.** (1993) Identification of control elements 3' to the human keratin-1 gene that regulate cell-type and differentiation-specific expression. *J. Biol. Chem.* 268: 377-384
- Hung, M. and Wensink, P.C.** (1981) The sequence of the *Drosophila melanogaster* gene for yolk protein 1. *Nucl. Acids Res.* 9: 6407-6419
- Hung, M. and Wensink, P.C.** (1983) Sequence and structure conservation in yolk proteins and their genes. *J. Mol. Biol.* 164: 481-489
- Hung, M., Barnett, T., Woolford, C. and Wensink, P.C.** (1982) Transcript maps of *Drosophila* yolk protein genes. *J. Mol. Biol.* 154: 581-602

- Huttner, W.B. and Baeuerle, P.A.** (1988) Protein sulfation on tyrosine. *Modern Cell Biology* 6: 97-140
- Ingolia, T.D., Craig, E.A. and McCarthy, B.J.** (1980) Sequence of three copies of the gene for the major *Drosophila* heat shock induced protein and their flanking regions. *Cell* 21: 669-679
- Inoue, K., Hoshijima, K., Higuchi, I., Sakamoto, H. and Shimura, Y.** (1992) . Binding of the *Drosophila transformer* and *transformer-2* proteins to the regulatory elements of *doublesex* primary transcript for sex-specific RNA processing. *Proc. Natl. Acad. Sci. USA* 89: 8092-8096
- Ireland, R.C. and Berger, E.M.** (1982) Synthesis of low molecular weight heat shock peptides stimulated by ecdysterone in a cultured *Drosophila* cell line. *Proc. Natl. Acad. Sci. USA* 79: 855-859
- Isaac, P.G.** (1982) Molecular studies of the vitellogenins of *Drosophila melanogaster*. Ph.D Thesis, University of Edinburgh.
- Isaac, P.G. and Bownes, M.** (1982) Ovarian and fat body vitellogenin synthesis in *Drosophila melanogaster*. *Eur. J. Biochem.* 123: 527-534
- Jacob, J. and Stirling, J.L.** (1959) Cell function in the ovary of *Drosophila* I. DNA classes in nurse cell nuclei as determined by autoradiography. *Chromosoma* 10: 210-228
- Jacoby, D.B. and Wensink, P.C.** (1994) Yolk protein factor 1 is a *Drosophila* homolog of Ku, the DNA-binding subunit of a DNA-dependent protein kinase from humans. *J. Biol. Chem.* 269: 11484-11491
- Jin, J. and Petri, W.H.** (1993) Developmental control elements in the promoter of a *Drosophila* vitelline membrane gene. *Dev. Biol.* 156: 557-565
- Jowett, T. and Postlethwait, J.H.** (1980) The regulation of yolk polypeptide synthesis in *Drosophila* ovaries and fat body by 20-hydroxyecdysone and a juvenile hormone analog. *Dev. Biol.* 80: 225-234
- Jursnich, V.A. and Burtis, K.C.** (1993) A positive role in differentiation for the male *doublesex* protein of *Drosophila*. *Dev. Biol.* 155: 235-249

- Kadesh, T., Zervos, P. and Ruezinsky, D.** (1986) Functional analysis of the murine IgH enhancer: Evidence for negative control of cell-type specificity. *Nucleic Acids Res.* 20: 8209-8221
- Karim, F.D. and Thummel, C.S.** (1991) Ecdysone coordinates the timing and amounts of E74A and E74B transcription in *Drosophila*. *Genes Dev.* 5:1067-1079
- Karim, F.D. and Thummel, C.S.** (1992) Temporal coordination of regulatory gene expression by the steroid hormone ecdysone. *EMBO J.* 11: 4083-4093
- Keyes, L.N., Cline, T.W. and Schedl, P.** (1992) The primary sex determination signal of *Drosophila* acts at the level of transcription. *Cell* 68: 933-943
- King, R.C.** (1970) Ovarian development of *Drosophila melanogaster*. Academic Press, New York
- Koelle, M.R., Talbot, S.W., Segraves, W.A., Bender, M.T., Cherbas, P. and Hogness, D.S.** (1991) The *Drosophila* EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell* 67: 59-77
- Kollias, G., Hurst, J., de Boer, E. and Grosveld, F.** (1987) The human  $\beta$ -globin gene contains a downstream developmental specific enhancer. *Nuc. Acids Res.* 15: 5739-5747
- Krimpenfort, P., de Jong, R., Uematsu, Y., Dembic, Z., Ryser, S., van Boehmer, H., Steinmetz, M. and Bernz, A.** (1988) Transcription of T cell receptor  $\beta$ -chain is controlled by a downstream regulatory element. *EMBO J.* 7: 745-750
- Laemmli, U.K.** (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T<sub>4</sub>. *Nature* 227: 680-685
- Laimins, L., Holmgren-Konig, M. and Khoury, G.** (1986) Transcriptional "silencer" element in rat repetitive sequences associated with the rat insulin I gene locus. *Proc. Natl. Acad. Sci. USA* 83: 3151-3155
- Langueux, M., Hetru, C., Goltzene, F., Kappler, C. and Hoffman, J.A.** (1979) Ecdysone titer and metabolism in relation to cuticulogenesis in embryos of *Locusta migratoria*. *J. Insect Physiol.* 25: 709-723



- Larsen, P.R., Harney, J.W. and Moore, D.D.** (1986) Repression mediates cell-type-specific expression of the rat growth hormone gene. *Proc. Natl. Acad. Sci. USA* 83: 8283-8287
- Lasko, P.F.** (1994) Molecular genetics of *Drosophila* oogenesis. M.B.I.U. R.G Laudes Co., Austin. CRC Press
- Laudet, V., Hänni, C., Coll, J., Catzeflis, F. and Stéhelin, D.** (1992) Evolution of the nuclear receptor gene superfamily. *EMBO J.* 11: 1003-1013
- Lauster, R., Reynaud, C-A., Mårtensson, I-L., Peter, A., Bucchini, D., Jami, J. and Weill, J-C.** (1993) Promoter, enhancer and silencer elements regulate rearrangement of an immunoglobulin transgene. *EMBO J.* 12: 4615-4623
- Liddell, S.** (1990) Expression of the yolk protein 3 gene of *Drosophila melanogaster*. Ph.D Thesis, University of Edinburgh
- Liddell, S. and Bownes, M.** (1991) Investigation of *cis*-acting sequences regulating expression of the gene encoding yolk protein 3 in *Drosophila melanogaster*. *Mol. Gen. Genet.* 230: 219-224
- Lindsley, D.L. and Grell, E.N.** (1968) Genetic variations of *Drosophila melanogaster*. Carnegie Institute of Washington, Publication no. 627
- Logan, S.K., Garabedian, M.J. and Wensink, P.C.** (1989) DNA regions that regulate the ovarian transcriptional specificity of *Drosophila* yolk protein genes. *Genes Dev.* 3: 1453-1461
- Logan, S.K. and Wensink, P.C.** (1990) Ovarian follicle cell enhancers from the *Drosophila* yolk protein genes: different segments of one enhancer have different cell-type specificities that interact to give normal expression. *Genes Dev.* 3: 613-623
- Lüning, K.G.** (1981) Genetics of inbred *Drosophila melanogaster*. *Hereditas* 95: 181-188
- Mahowald, A.P. and Strassheim, J.M.** (1970) Intercellular migration of centrioles in the germarium of *Drosophila melanogaster*. An electron microscopic study. *J. Cell. Biol.* 45: 306-320

- Mahowald, A.P. and Tiefert, M.** (1970) Some structural changes in the *Drosophila* oocyte nucleus during a short period of RNA synthesis. Wilhelm Roux Arch. Entw. Org. 165: 8-25
- Mahowald, A.P. and Kambyzellis, M.P.** (1980) Oogenesis. In: The genetics and biology of *Drosophila*, vol 2. Eds. Ashburner, M. and Wright, TRF. Academic Press, New York, pp 141-224
- Maniatis, T., Fritsch, E.F. and Sambrook, J.** (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbour Laboratories, Cold Spring Harbour, New York.
- Marcus, C.H.** (1985) Dros. Info. Serv. 61: 193-194
- Margaritis, L.H., Kafatos, F.C. and Petri, W.H.** (1980) The egg shell of *Drosophila melanogaster*, I. Fine structure of the layers and regions of the wild-type eggshell. J. Cell Sci. 43: 1-35
- Mariani, B.D., Lingappa, J.R. and Kafatos, F.C.** (1988) Temporal regulation in development: Negative and positive cis-regulators dictate the precise timing of expression of a *Drosophila* chorion gene. Proc. Natl. Acad. Sci. USA 85: 3029-3033
- Martinez, A. and Bownes, M.** (1992) The specificity of yolk protein uptake in Cyclorrhaphan Diptera is conserved through evolution. J. Mol. Evol. 35: 444-453
- Martinez, A. and Bownes, M.** (1994) The sequence and expression pattern of the *Calliphora erythrocephala* yolk protein A and B genes. J. Mol. Evol. 38: 336-351
- Martinez, E., Givel, F. and Wahli, W.** (1991) A common ancestor DNA motif for invertebrate and vertebrate hormone response elements. EMBO J. 10: 263-268
- Matthias, P. and Baltimore, D.** (1993) The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the  $\alpha$  constant region. Mol. Cell. Biol. 13: 1547-1553
- Melton, D.A., Kreig, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.K.** (1984) Efficient *in vitro* synthesis of biologically active RNA and RNA hybridisation probes from plasmids containing a bacteriophage SP6 promoter. Nucl. Acids Res. 12: 7035-7056

- Messing, J.** (1983) New M13 vectors for M13mp19 cloning. *Meths. Enzymol.* 101: 20-78
- Meyer, K.B. and Neuberger, M.S.** (1989) The immunoglobulin kappa-locus contains a second, stronger, B-cell-specific enhancer which is located downstream of the constant region. *EMBO J.* 8: 1959-1964
- Meyer, M-E., Gronemeyer, H., Turcotte, B., Bocquel, M-T., Tasset, D. and Chambon, P.** (1989) Steroid hormone receptors compete for factors that mediate their enhancer function. *Cell* 57: 433-442
- Mindrinos, M.N., Scherer, L.J., Garcini, F.J., Kwan, H., Jacobs, K.A. and Petri, W.H.** (1985) Isolation and chromosomal location of putative vitelline membrane genes in *Drosophila melanogaster*. *EMBO J.* 4: 147-153
- Minoo, P. and Postlethwait, J.** (1985) Biosynthesis of *Drosophila* yolk polypeptides. *Archives of Insect Biochem. and Physiol.* 2: 7-27
- Mintzas, A. and Kambysellis, M.P.** (1982) The yolk proteins of *Drosophila melanogaster*: isolation and characterisation. *Insect Biochem.* 12: 25-33
- Mitsis, P.G. and Wensink, P.C.** (1989a) Identification of yolk protein factor 1, a sequence-specific DNA-binding protein from *Drosophila melanogaster*. *J. Biol. Chem.* 264: 5188-5194
- Mitsis, P.G. and Wensink, P.C.** (1989b) Purification and properties of yolk protein factor 1, a sequence-specific DNA-binding protein from *Drosophila melanogaster*. *J. Biol. Chem.* 264: 5195-5202
- Montell, D.J., Rorth, P. and Spradling, A.C.** (1992) *slow border cells*, a locus required for a developmentally regulated cell migration during oogenesis, encodes *Drosophila* C/EBP. *Cell* 71: 51-62
- Morganelli, C.M., Berger, E.M. and Pelham, H.R.B.** (1985) Transcription of *Drosophila* small hsp-tk hybrid genes is induced by heat shock and by ecdysterone in transfected *Drosophila* cells. *Proc. Natl. Acad. Sci. USA* 82: 5865-5869
- Müller, J. and Bienz, M.** (1991) Long range repression conferring boundaries of *ultrabithorax* expression in the *Drosophila* embryo. *EMBO J.* 10: 3147-3155

- Oliver, B., Perrimon, N. and Mahowald, A.** (1988) Genetic evidence that the *sans-fille* locus is involved in *Drosophila* sex-determination. *Genetics* 120: 159-171
- Orr-Weaver, T.L.** (1991) *Drosophila* chorion genes : cracking the eggshell's secrets. *BioEssays* 13: 97-105
- Osir, E.O., Anderson, D.R., Grimes, W.J. and Law, J.H.** (1986) Studies on the carbohydrate moiety of vitellogenin from the tobacco hornworm, *Manduca sexta*. *Insect Biochem.* 16: 471-478
- Palli, S.R., Touhara, K., Charles, J-P., Bonning, B.C., Atkinson, J.K., Trowell, S.C., Hiruma, K., Goodman, W.S., Kyriakides, T., Prestwich, G.D., Hammock, B.D. and Riddiford, L.M.** (1994) A nuclear juvenile hormone-binding protein from larvae of *Manduca sexta*: a putative receptor for the metamorphic action of JH. *Proc. Natl. Acad. Sci. USA* 91: 6191-6195
- Parkhurst, S.M., Bopp, D. and Ish-Horowics, D.** (1990) X:A ratio, the primary sex-determining signal in *Drosophila*, is transduced by helix-loop-helix proteins. *Cell* 63:1179-1191
- Parks, S., Wakimoto, B. and Spradling, A.C.** (1986) Replication and expression of an X-linked cluster of *Drosophila* chorion genes. *Dev. Biol.* 117: 294-305
- Pelham, H.R.B.** (1982) A regulatory upstream promoter element in the *Drosophila* hsp70 heat-shock gene. *Cell* 30: 517-528
- Pelham, H.R.B. and Bienz, M.** (1982) A synthetic heat-shock promoter element confers heat-inducibility on the herpes simplex virus thymidine kinase gene. *EMBO J.* 1: 1473-1477
- Perez-Stable, C. and Constantini, F.** (1990) Roles of fetal  $\gamma$ -globin promoter elements and the adult  $\beta$ -globin 3' enhancer in the stage-specific expression of globin genes. *Mol. Cell. Biol.* 10: 1116-1125
- Persson, B., Bergtsson-Olivecrona, G., Enerback, S., Olivecrona, T. and Jornvall, H.** (1989) Structural features of lipoprotein lipase. *Eur. J. Biochem.* 179: 39-45
- Postlethwait, J.H. and Handler, A.M.** (1978) Non-vitellogenic female sterile mutants and the regulation of vitellogenesis in *Drosophila melanogaster*. *Dev. Biol.* 67: 202-213

- Postlethwait, J.H. and Handler, A.M.** (1979) The roles of juvenile hormone and 20-hydroxyecdysone during vitellogenesis in isolated abdomens of *Drosophila melanogaster*. *J. Insect Physiol.* 25: 455-460
- Postlethwait, J.H. and Jowett, T.** (1980) Genetic analysis of the hormonally regulated yolk polypeptide genes in *D. melanogaster*. *Cell* 20: 671-678
- Postlethwait, J.H. and Weiser, K.** (1973) Vitellogenesis induced by juvenile hormone in the female sterile mutant *apterous-four* in *Drosophila melanogaster*. *Nature (London)* 244: 284-285
- Postlethwait, J.H., Bownes, M. and Jowett, T.** (1980) Sexual phenotype and vitellogenin synthesis in *Drosophila melanogaster*. *Dev. Biol.* 79: 378-387
- Query, C.C., Bentley, R.C and Keene, J.D.** (1989) A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell* 57: 89-101
- Racciopi, J.V., Gemmill, R.M., Kogan, P.H., Catto, J.M. and Hagedorn, H.H.** (1986) Expression and regulation of vitellogenin messenger RNA in the mosquito *Aedes aegypti*. *Insect Biochem.* 16: 255-262
- Read, S.M. and Northcote, D.H.** (1981) Minimisation of variation in the response to different proteins of the Coomassie blue G dye binding assay of protein. *Anal. Biochem.* 116: 53-64
- Redfern, C.P.F. and Bownes, M.** (1982) Yolk polypeptide synthesis and uptake in ovaries of the *apterous-4* mutant of *Drosophila melanogaster*. *Mol. Gen. Genet.* 185: 181-185
- Riddough, G. and Pelham, H.R.B.** (1987) An ecdysone response element in the *Drosophila* hsp27 promoter. *EMBO J.* 6: 3729-3734
- Rohrkasten, A. and Ferenz, H.J.** (1986) Properties of the vitellogenin receptor of isolated locust oocyte membranes. *Int. J. Inv. Rep. Dev.* 10: 133-142
- Ronaldson, E. and Bownes, M.** (1995) Two independent cis-acting elements regulate the sex- and tissue-specific expression of *yp3* in *Drosophila melanogaster*. *Genet. Res.* In press
- Roseman, R.R., Pirrotta, V. and Geyer, P.K.** The *su(Hw)* protein insulates expression of the *Drosophila melanogaster* *white* gene from chromosomal position-effects. *EMBO J.* 12: 435-442

- Rubin, G.M. and Spradling, A.C.** (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science* 218: 348-353
- Ryner, L.C. and Baker, B.S.** (1991) Regulation of *doublesex* pre-mRNA processing occurs by 3'-splice site activation. *Genes Dev.* 5: 2071-2085
- Salz, H.K.** (1992) The genetic analysis of *snf*: A *Drosophila* sex determination gene required for activation of *Sex-lethal* in both the germline and the soma. *Genetics* 130: 547-554
- Savagner, P., Miyashita, T. and Yamada, Y.** (1990) Two silencers regulate the tissue-specific expression of the collagen II gene. *J. Biol. Chem.* 265: 6669-6674
- Savakis, C., Demetri, G. and Cherbas, P.** (1980) Ecdysteroid-inducible polypeptides in a *Drosophila* cell line. *Cell* 22: 665-674
- Savakis, C., Koehler, M.M.D. and Cherbas, P.** (1984) cDNA clones for the ecdysone-inducible polypeptide (EIP) mRNAs of *Drosophila* Kc cells. *EMBO J.* 3: 235-243
- Savant, S.S. and Waring, G.L.** (1989) Molecular analysis and rescue of a vitelline membrane mutant in *Drosophila*. *Dev. Biol.* 135: 43-52
- Schwartz, M.B., Kelly, T.J., Imberski, R.B. and Rubenstein, E.C.** (1985) The effects of nutrition and methoprene treatment on ovarian ecdysteroid synthesis in *Drosophila melanogaster*. *J. Insect Physiol.* 31: 947-959
- Segraves, W.A. and Hogness, D.S.** (1990) The E75 ecdysone-inducible gene responsible for the 75B early puff in *Drosophila* encodes two new members of the steroid receptor superfamily. *Genes Dev.* 4: 204-219
- Serfling, E., Jasin, M. and Schaffner, W.** (1985) Enhancers and eukaryotic gene transcription. *Trends Genet.* 1: 224-230
- Shea, M.J., King, D.L., Conboy, M.J., Mariani, B.D. and Kafatos, F.C.** (1990) Proteins that bind to *Drosophila* chorion *cis*-regulatory elements: A new C<sub>2</sub>H<sub>2</sub> zinc finger protein and a C<sub>2</sub>C<sub>2</sub> steroid receptor-like component. *Genes Dev.* 4: 1128-1140
- Shirras, A. and Bownes, M.** (1987) Separate DNA sequences are required for normal female and ecdysone-induced male expression of *Drosophila melanogaster* yolk protein 1. *Mol. Gen. Genet.* 210: 153-155

- Shirras, A. and Bownes, M.** (1989) *cricket*: A locus regulating a number of adult functions of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 86: 4559-4563
- Singh, H.** (1993) Meths. Enzymol. 218: 551-567
- Slee, R. and Bownes, M.** (1990) Sex determination in *Drosophila melanogaster*. Quarterly Review of Biology 65: 175-204
- Spradling, A.C. and Rubin, G.M.** (1982) Transposition of cloned P elements into *Drosophila* germline chromosomes. Science 218: 341-347
- Steinmann-Zwicky, M.** (1988) Sex-determination in *Drosophila*: the X chromosome gene *liz* is required for *Sxl* activity. EMBO J. 7: 3889-3898
- Steinmann-Zwicky, M., Amrein, H. and Nöthiger, R.** (1990) Genetic control of sex determination in *Drosophila*. Adv. Genet. 27: 189-237
- Takemoto, Y., Fujimura, Y., Matsumoto, M., Tamai, Y., Morita, T., Matsushiro, A. and Nozaki, M.** (1991) The promoter of the *endo A* cytochrome gene is activated by a 3' downstream enhancer. Nuc. Acids Res. 19: 2761-2765
- Talbot, W.S., Swyryd, E.A. and Hogness, D.S.** (1993) *Drosophila* tissues with different metamorphic responses to ecdysone express different ecdysone receptor isoforms. Cell 73: 1323-1337
- Tamura, T., Kunert, C. and Postlethwait, J.** (1985) Sex- and cell-specific regulation of yolk polypeptide genes introduced into *Drosophila* by P-element-mediated germ transfer. Proc. Natl. Acad. Sci. USA 82: 7000-7004
- Thomas, H.E., Stunneberg, H.G. and Stewart, A.F.** (1993) Heterodimerization of the *Drosophila* ecdysone receptor with retinoid X receptor and *ultraspiracle*. Nature 362: 471-475
- Thummel, C.S., Boulet, A.M. and Lipshitz, H.D.** (1988) Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. Gene 74: 445-456
- Thummel, C.S., Burtis, K.C. and Hogness, D.S.** (1990) Spatial and temporal patterns of *E74* transcription during *Drosophila* development. Cell 61: 101-111

- Tolias, P.P. and Kafatos, F.C.** (1990) Functional dissection of an early *Drosophila* chorion gene promoter: expression throughout the follicular epithelium is under spatially composite regulation. *EMBO J.* 9: 1457-1464
- Torres, M. and Sánchez, L.** (1989) The *scute* (*T4*) gene acts as a numerator element of the X:A signal that determines the state of activity of *Sex-lethal* in *Drosophila*. *EMBO J.* 8: 3079-3086
- Torres, M. and Sánchez, L.** (1992) The segmentation gene *run* is needed to activate *Sex-lethal*, a gene that controls sex determination and dosage compensation in *Drosophila*. *Genet. Res.* 59: 189-198
- van Doren, M., Ellis, H.M. and Posakony, J.W.** (1991) The *Drosophila extramacrochaetae* protein antagonizes sequence-specific DNA binding by *daughterless/achaeta-scute* protein complexes. *Development* 113: 245-255
- van Handel, E. and Lea, A.O.** (1984) Vitellogenin synthesis in blood-fed *Aedes aegypti* in the absence of the head, thorax and ovaries. *J. Insect Physiol.* 30: 871-875
- Wada-Kiyama, Y., Peters, B. and Noguchi, C.T.** (1992) The  $\epsilon$ -globin gene silencer. *J. Biol. Chem.* 267: 11532-11538
- Warren, T.G. and Mahowald, A.P.** (1979) Isolation and partial chemical characterization of the three major yolk polypeptides from *Drosophila melanogaster*. *Dev. Biol.* 68: 130-139
- Williams, J.L. and Bownes, M.** (1986) Reduced stability of RNA coding for yolk polypeptide 3 in *Drosophila melanogaster* ovary. *Eur. J. Biochem.* 161: 95-101
- Winoto, A. and Baltimore, D.** (1989) A novel inducible and T-cell-specific enhancer located at the 3' end of the T-cell receptor alpha locus. *EMBO J.* 3: 729-733
- Wolfner, M.F.** (1988) Sex-specific gene expression in somatic tissues of *Drosophila melanogaster*. *Trends Genet.* 4: 333-337
- Wu, S., Zhang, J. and Ma, M.** (1987) Monitoring the effects of juvenile hormones and 20-hydroxyecdysone on yolk polypeptide production of *Drosophila melanogaster* with enzyme immunoassay. *Physiol. Ent.* 12: 355-361
- Wyatt, G.R.** (1991) Gene regulation in insect reproduction. *Invert. Reprod. Dev.* 20: 1-35



- Xiao, H. and Lis, J.T.** (1988) Germline transformation used to define key features of heat-shock response elements. *Science* 239: 1139-1142
- Yan, Y.L., Kunert, C. and Postlethwait, J.H.** (1987) Sequence homologies among the three yolk polypeptide (Yp) genes in *Drosophila melanogaster*. *Nuc. Acids Res.* 15: 67-85
- Yao, T-P., Segraves, W.A., Oro, A.E., McKeown, M. and Evans, R.M.** (1992) *Drosophila* ultraspiracle modulates ecdysone receptor function via heterodimer formation. *Cell* 71: 63-72
- Younger-Shepherd, S., Vaessin, H., Bier, E., Jan, L.Y. and Jan, Y.N.** (1992) *deadpan*, an essential pan-neural gene encoding an HLH protein, acts a denominator in *Drosophila* sex determination. *Cell* 70: 911-922